Exhibit A

Chapter 6: Common Device Elements

B. Stark

While a completed MEMS is a complicated device, the individual components of any given system are much simpler to understand. Due to the nature of MEMS processing, no single component can be very complex. This in turn means that understanding of a MEMS device can be gained through knowledge of a few simple parts and understanding how they interact. To ensure the reliable operation of a MEMS device it is sufficient to ensure the reliable operation of all the constituent parts.

One of the difficulties in writing a guideline is trying to select material that will not be dated before the book goes to press. To preclude this problem, this chapter does not address specific sensor technologies, but rather it deals with device elements. It is assumed that a knowledgeable reliability engineer can construe all the necessary information on, for example, a capacitive accelerometer by examining the sections on structural beams and parallel plate capacitors. It is felt that this arrangement of the material will increase its useful lifetime.

This chapter has been loosely organized into three sections. The first three subchapters discuss structural elements in MEMS. The next two subchapters, along with part of the third, discuss transducer elements. The remainder of the chapter is dedicated to actuator technologies.

I. Structural Beams

Structural beams are a basic building block of most MEMS devices. A beam is, as the name implies, a long thin piece of material that often serves as the supporting basis for a structure. While this is fairly self-evident, there are also some commonalities specific to MEMS beams that are useful to understand. The majority of beams used in MEMS have the rectangular cross section defined below:

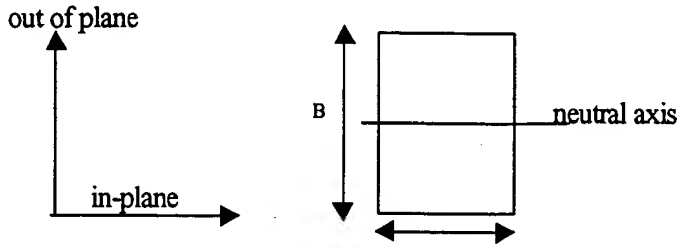


Figure 6-1: Cross section of common MEMS beams. Dimension A is planar and is limited by the minimum feature size of the processing technology. Dimension B is non planar and is limited by the aspect ratio of the etching technology.

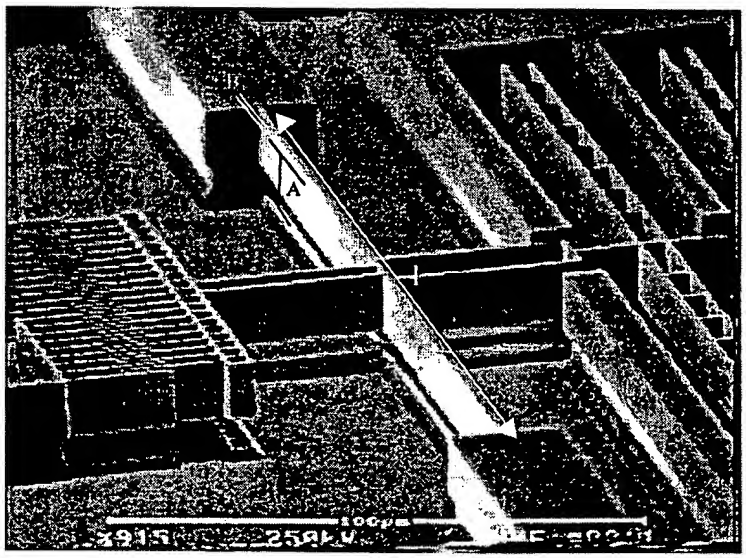


Figure 6-2: SEM picture of beam with all dimensions labeled. (from [155])

A SEM picture of a beam is included in Figure 6-2 to show the usual labels given to each beam dimension. It should also be noted that the rectangular cross section is a generally

accepted approximation, as there are multiple non-uniform features introduced in device construction. In some cases, such as beams made with wet etching and certain CMOS processes, the cross section is more trapezoidal than rectangular. However, once the basic shape of the beam is determined, it is relatively simple to construe its mechanical properties.

A. Structural Analysis of Support Beams

i) Static Deflections

Support beams are analyzed for both reliability and performance using techniques common to most engineering students. One issue critical to understanding beams is understanding how they bend under different loadings. The most common method to determine this involves the Euler-Bernoulli equation:

$$\frac{d^2y}{dx^2} = \frac{M(x)}{EI} \tag{6-1}$$

where

x = direction along the neutral axis

y = direction along the transverse axis

E = Young's modulus

I = area moment of inertia

M(x) = the bending moment in the beam, which is usually a function of x.

As this definition may be a bit obtuse, Figure 6-3 illustrates the concept being described.

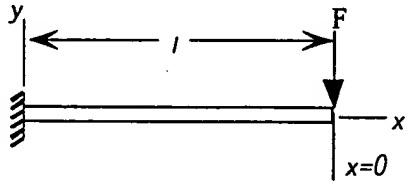
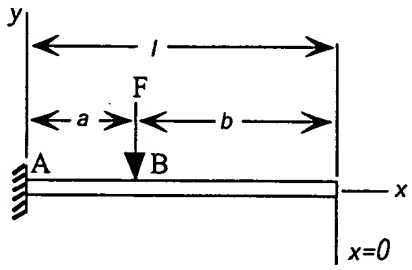
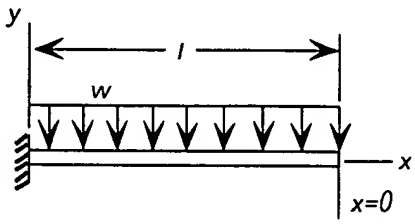
Beam Configuration	Vertical (y) Displacement
	$y(x) = -\frac{F}{EI}\left(\frac{x^3}{6} - \frac{l^2 x}{2} + \frac{l^3}{3}\right)$ $y_{max} = -\frac{Fl^3}{3EI}$
	$y_{AB} = -\frac{F}{EI}\left(\frac{(x-b)^3}{6} - \frac{a^2 x}{2} + \left(\frac{la^2}{2} - \frac{a^3}{6}\right)\right)$ $y_{BC} = -\frac{F(b-x)}{EI}\left(\frac{a^2}{2}\right) - \frac{F}{EI}\left(\frac{ba^2}{2} + \frac{la^2}{2} - \frac{a^3}{6}\right)$ $y_{max} = -\frac{Fb}{EI}\left(\frac{a^2}{2}\right) - \frac{F}{EI}\left(\frac{ba^2}{2} + \frac{la^2}{2} - \frac{a^3}{6}\right)$
	$y = -\frac{w}{24EI}(x^4 - 4l^3 x + 3l^4)$ $y_{max} = -\frac{wl^4}{8EI}$

Figure 6-3: Displacement of loaded cantilever beams.

To analyze the deformation of a beam under transverse loading, Equation 6-1 is integrated twice using the appropriate boundary conditions. This yields the result:

$$y(x) = -\frac{1}{EI}\iint M(x)dxdx + Cx + D \qquad (6\text{-}2a)$$

$$y'(x) = -\frac{1}{EI}\int M(x)dx + C \tag{6-2b}$$

where C and D are constants determined by the boundary conditions.

For a cantilever beam, which is one of the most structural beams in MEMS, with the boundary conditions of y(l)=0 and y'(l)=0 and a force, F, applied at one end, the equation yields:

$$y(x) = -\frac{F}{EI}\left(\frac{x^3}{6} - \frac{l^2 x}{2} + \frac{l^3}{3}\right) \tag{6-3}$$

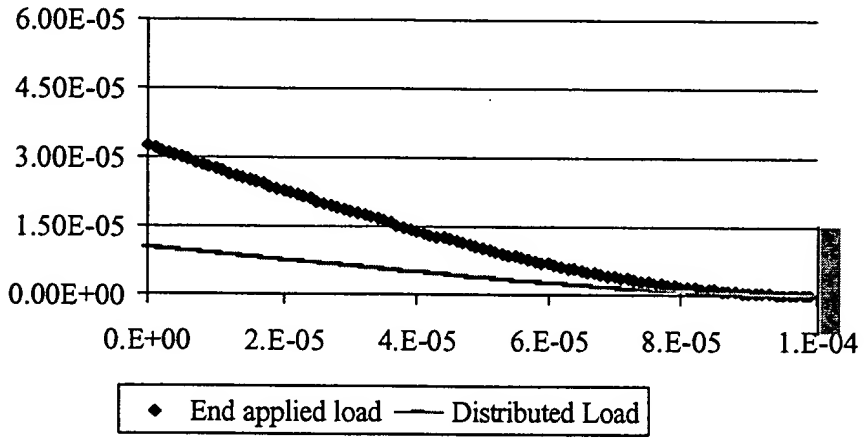


Figure 6-4: Graphical representation of a cantilever beam deforming under a transverse load.

Figure 6-3 contains the equations describing cantilever deflections under other common loading conditions. This analysis also leads to another piece of information that is useful to understand. Since Equation 6-3 describes a linear force-deflection relationship, it is essentially describing a spring reacting to an applied load. This means that it is possible to extract a spring constant, k, from this expression. Evaluating y(x) at a specific point will determine the spring constant. For this example y(0), which is equal to $-Fl^3/3EI$, will be used. Rearranging this equation yields:

$$\frac{F}{y(0)} = -\frac{3EI}{l^3} \tag{6-4}$$

and, since F/y = k, this gives the result:

$$k_y = \frac{3EI}{l^3} \tag{6-5a}$$

The value of I can either be determined by integration or by tables. For rectangular cross sections and planar bending, I is $a^3b/12$ and Equation 6-5a is rewritten as 6-5b. If a non-homogenous beam is used, then the method introduced in Chapter 5-2C describes how to normalize this beam to a uniform cross section.

$$k_y = \frac{Ea^3b}{4l^3} \tag{6-5b}$$

While this expression is useful for predicting displacement under a given load, there are some limitations to it that must be understood. Hooke's law of $F_x=kx$ only applies for small displacements. For larger displacements, non-linear terms will appear in the force-displacement equations. The degree to which this equation applies thus depends largely upon how large a force is applied to the structure. Often, to simplify the development of devices, designers will construct structures that will operate solely within the linear regime. However it is important to understand that the linear force-displacement equation is only a first order approximation of the actual relationship between force and displacement.

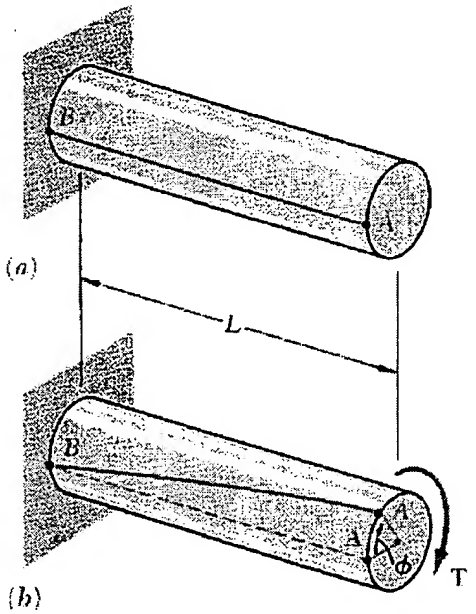


Figure 6-5: Beams displacing an angle, ϕ, due to an applied torque. (from [11])

For some applications, especially those involving non-planar displacements, it is necessary to subject structural beams to a torque. There are methods for analyzing rigid bodies under a torsional load, an example of which is shown in Figure 6-5. For these applications, it is useful to relate a torque to the angle of twist, ϕ, which can be accomplished with the expression:

$$\phi = \frac{TL}{c_2 G a^3 b} \tag{6-6}$$

where

T = applied torque

c_2 = a constant defined in Table 6-1

b/a	c_1	c_2
1.0	0.208	0.1406
1.5	0.231	0.1958
2.0	0.246	0.229
2.5	0.258	0.249
3.0	0.267	0.263
4.0	0.282	0.281
5.0	0.291	0.291
10.0	0.312	0.312
∞	0.333	0.333

Table 6-1: Constants for rectangular cross sections under a torsional load.[11]

As might be expected, there is also a torsional stiffness related to beam geometry that is useful in analyzing non planar actuators. This stiffness, k_ϕ is defined by:

$$k_\phi \equiv \frac{T}{\phi} = 2c_2 G \frac{a^3 b}{l} \tag{6-7}$$

ii) Oscillatory Motion

As structural beams are often operated in resonant modes, it is necessary to analyze the oscillatory motion of beams. Resonant frequency[1] is determined by the equation:

[1] The terms "resonant frequency" and "natural frequency" are used interchangeably in this section. Although this is common in the literature, these are actually two distinct quantities. The relationship between resonant and natural frequency is discussed in detail in Section 3-VII.

$$\omega_0 = \sqrt{\frac{k}{m_{eff}}} \tag{6-8}$$

where

k = stiffness or spring constant

m_{eff} = moving mass

The only two quantities required to determine the natural frequency of a beam are k and m_{eff}. Since k has already been derived, it is necessary to calculate the moving mass. The moving mass can be analytically determined using Rayleigh's method. However, this method exceeds the scope of the guideline. It suffices to know that, for a cantilever beam, the moving mass is roughly 23% of the total mass. This can be analytically described by

$$m_{eff} = .23\rho abl \tag{6-9}$$

where ρ is the mass density of the beam. This leads to a final expression for resonant frequency:

$$\omega_0 = 1.043\frac{a}{l^2}\sqrt{\frac{E}{\rho}} \tag{6-10}$$

In cases when beams oscillate in torsion, the torsional resonant frequency is:

$$\omega_0 = \sqrt{\frac{k_\phi}{I_\phi}} \tag{6-11}$$

Since beams driven by harmonic transverse loads behave similarly to strings in tension, the wave equation can describe analytically how a beam moves in resonance:

$$\frac{\partial F}{\partial x} = -\frac{m}{l}\frac{\partial^2 y}{\partial t^2} \tag{6-12a}$$

Through a substitution of the force for an expression involving the moment, this equation becomes:

$$-\frac{\partial^2}{\partial x^2}\left(EI\frac{\partial^2 y}{\partial x^2}\right) = \frac{m}{l}\frac{\partial^2 y}{\partial t^2} \tag{6-12b}$$

This equation describes the curvature of a beam as a function of time. Through the study of differential equations, Equation 6-12b is solved as:

$$y = (A_1 \sin(Kx) + A_2 \cos(Kx) + A_3 \sinh(Kx) + A_4 \cosh(Kx))\cos(\omega t + \Theta) \qquad (6\text{-}13)$$

where

A_n = constant determined by the boundary conditions

$$K^4 = \frac{\omega_n^2 m}{EI}$$

This reveals that, at resonance, a beam will oscillate in a sinusoidal fashion, with the shape of the beam determined by the boundary conditions:

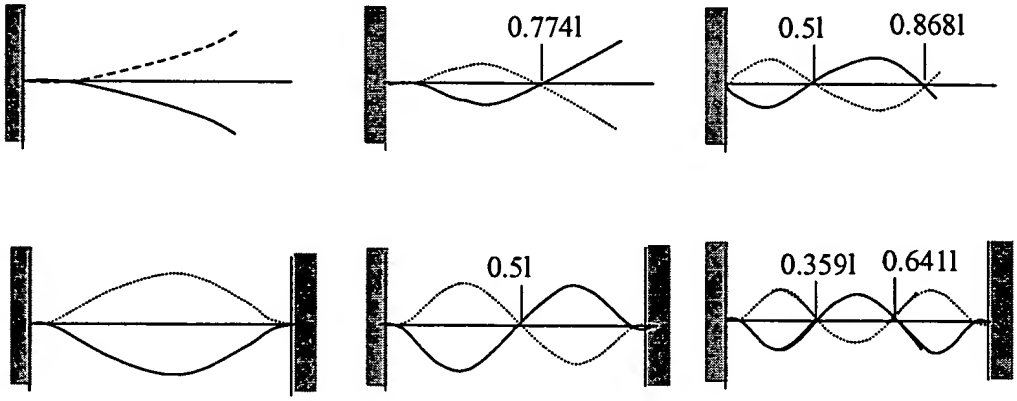


Figure 6-6: Oscillatory modes for cantilever beams (top) and built-in beams (bottom).

While using the above techniques will determine the shape of beams under a variety of loads, different methods will have to be introduced to insure that the beams will not fail under the stresses caused by these loads.

B. Failure of Structural Beams

A structural beam will fail when the maximum allowable stress has been exceeded. For different types of materials, different failure modes are exhibited, as discussed in Chapter 3. However, for reliable operation, the stresses in materials should not approach either yield or ultimate stress. There are several methods useful for calculating stresses in materials.

The normal stress, σ_x, is easy to calculate. It can be determined by the relationship:

$$\sigma_x = \frac{My}{I} \tag{6-14}$$

where y is the vertical distance from the neutral axis.

While this equation is valid for small deformations, for large deformations, the slope of a beam at the loading point[1], $\left.\frac{dy}{dx}\right|_l$, becomes important. For large deflections the maximum stress is:

$$\sigma_{\max}^{large} = \frac{y}{I}\sqrt{\frac{2EIM\left.\frac{dy}{dx}\right|_l}{I(1-\nu^2)}} \tag{6-15}$$

The ratio of $\sigma_{\max}^{small} / \sigma_{\max}^{large}$ is compared to get an estimate in the error in using the linearized approximation of Equation 6-14. One study [40] found the error to be 5% at $\tan^{-1}\left(\left.\frac{dy}{dx}\right|_l\right) = 30°$, 11% at 45° and 22% at 60°. So, the extent to which a linear approximation is valid depends upon the size of the deflection, as one would expect.

Determination of shearing stress is slightly more difficult. Shearing stress is calculated by examining a section of a beam and using the equation:

$$\tau_{ave} = \frac{VQ}{It} \tag{6-16}$$

where

τ_{ave} = average shearing stress in a section of the beam

V = vertical shear in a given cross section

t = thickness of the cross section

Q = 1st moment of the area defined as $Q = \int_{y=y1}^{y=C} ydA$

---

[1] This derivation assumes a concentrated load.

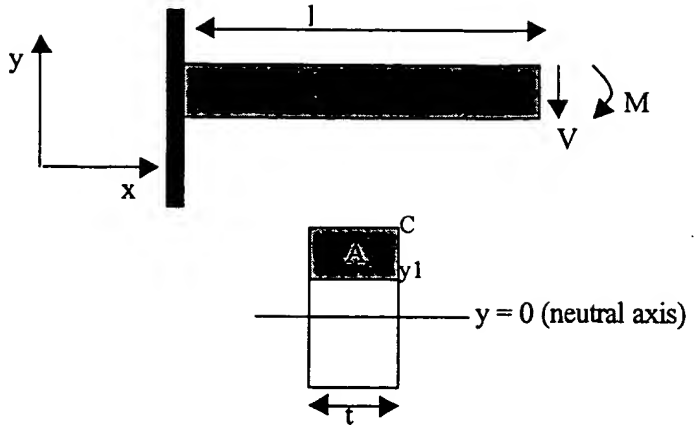


Figure 6-7: Side (top) and cross sectional (bottom) views of cantilever beam under shear force and bending movement.

The relationship among all of these values is presented in Figure 6-7. To determine τ_{max}, it is usually necessary to examine τ_{ave} as a function of y. For beams with the height > 4×width, the approximation of τ_{max}=1.5×V/(w×h) is useful. However, for planar bending in beams, where the b ≥ a relationship obviates this approximation, the full structural analysis will have to be performed. It is convenient to recognize that maximum shearing stress will always occur along the neutral axis, unless there are marked variations in beam thickness. Thus, determination of shearing stress in beams can usually be performed simply by examining τ_{ave} at y=0.

For rectangular beams subjected to a torque, there will be a shear stress that will vary as a function of horizontal and vertical position. However the study of mechanics reveals that the maximum stress, which is the most important in terms of structural analysis, occurs along the neutral axis of the wider face of the beam, which, in this discussion, is labeled as b. This maximum shearing stress is determined by the relation:

$$\tau_{max} = \frac{T}{c_1 a^2 b} \tag{6-17}$$

where c_1 is a constant defined in Table 6-1.

While using the above analysis leads to a good understanding of when fracture will occur, there are other failure modes in structural beams. Beams will usually fail if they come into contact with other structures, due to adhesive forces. To analyze the probability of this, the deflection of every beam must be considered under maximum load.

Another concern raised in Chapter 3 was the impact of fatigue in beams. Over long cycle times, the properties of beams will shift, which, as this section has shown, will change the static and resonant characteristics of the structures. These changes will alter the output of many sensors based upon measuring frequency and deflection. Another fatigue related mechanism is the gradual relaxation of the fracture strength of a material. Beams that were initially driven within stress tolerances of a material, can be driven past them, as the tolerances decrease. Although it should be noted again that this mechanism has yet to be observed in silicon.

Thermal changes can also have an impact upon beam reliability. Thermal stressing and unstressing creates mechanical fatigue in beams. In large temperature changes, as experienced in the space environment, most MEMS beams will also experience bimetallic warping due to the fact that they are made of different materials that have mismatched thermal coefficients of expansions. Thermal fatigue can also contribute to delamination.

C. Additional Reading

F. P. Beer and E. R. Johnston, Mechanics of Materials: Second Edition, McGraw-Hill. New York, 1992.

C. M. Harris and C. E. Crede, Shock and Vibration Handbook Volume 1: Basic Theory and Measurements, McGraw-Hill Book Company, Inc., New York, 1961.

II. Thin Membranes

In recent years, thin membranes have found increasing use in pressure and flow sensors. They provide a large sensing area coupled with low mass, which is advantageous in many applications. A membrane is commonly assumed to be any structure with a z dimension much smaller than its x and y dimensions. While the membranes used in MEMS do not fit the classic definition of plates, their thickness deformation can be influenced by in-plane tension. This term is also applied to these devices. A membrane structure is shown in Figure 6-8. The analysis of a membrane is more difficult than cantilever beams, but it is still tractable.

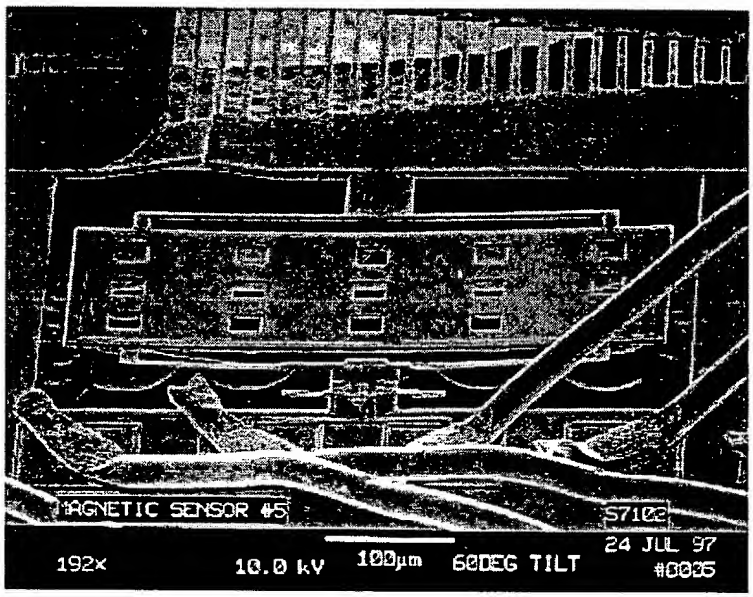


Figure 6-8: A thin plate viewed at 200x magnification. (from JPL)

A. Structural Analysis of Membranes

i) Static Deflection

There are several methods commonly employed in the analysis of thin plates or membranes. The most obvious method is to use the equations of motion to describe the plate, as was done for beams. This is accomplished by defining coordinates for the plate in the x and y axis and taking into account all shearing and bending forces. This analysis leads to a system of six equations and six unknowns which reduces to the result:

$$\nabla^4 w = \frac{q}{D} \tag{6-18}$$

where

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$$

w = the plate deflection at any given point

q = the lateral load function

D = the plate stiffness = $\frac{Eh^3}{12(1-\nu^2)}$

h = the plate thickness

The solution of this equation clearly requires determining the function w, such that it satisfies both the loading and boundary conditions. Since empirical data shows that these models are not the most accurate, the method developed by J.Y. Pan[14] is often used. This method begins by determining the midpoint deflection of a membrane, w_0. For a square membrane there is a relationship between midpoint deflection and an applied pressure p given by:

$$p(w_0) = C_1 \frac{h\sigma}{a^2} w_0 + C_2 \frac{hE}{a^4} w_0^3 \text{ [14]} \tag{6-19}$$

where

σ = internal stress

a = plate width

w_0 = plate deflection at center point

C_1, C_2 = functions of Poisson's ratio, defined in Figure 6-9a

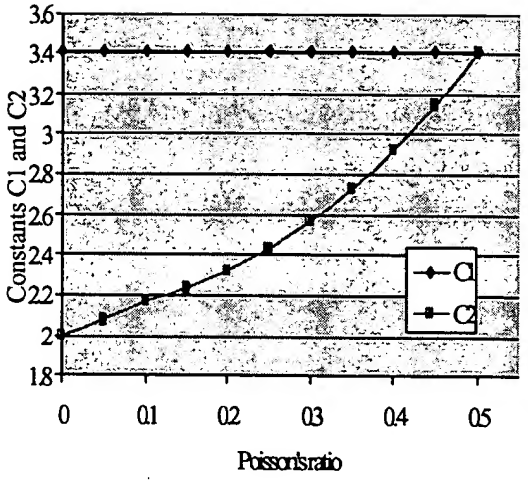


Figure 6-9a: Dependence of C_1 and C_2 upon Poisson's ratio.[15]

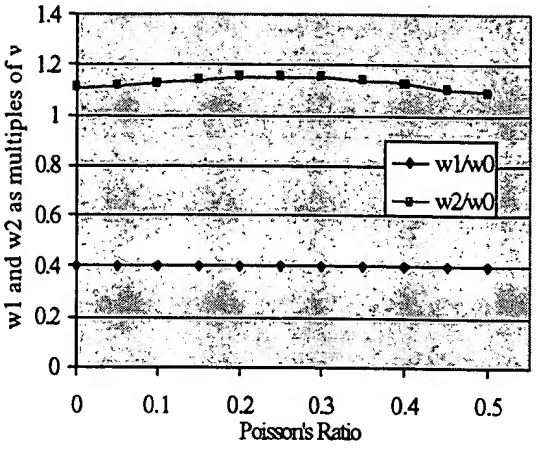


Figure 6-9b: w_1/w_0 and w_1/w_0 as a function of Poisson's ratio.[15]

Once w_0 is known, it is possible to determine the shape of the entire plate. If the origin of the plate is taken at its geometric center, then the deflection is described by:

$$w(x,y) = \left(w_0 + w_1 \frac{x^2 + y^2}{a^2} + w_2 \frac{x^2 y^2}{a^4} \right) \cos\left(\frac{\pi x}{2a}\right) \cos\left(\frac{\pi y}{2a}\right) \qquad (6\text{-}20)$$

where w_1 and w_2 are functions of Poisson's ratio related to w_0 by Figure 6-9b.

While this method has the attribute that it offers a closed form solution to the shape of the plate as a function of x and y, it is not always possible to solve the deformation of the plate so simply. In instances when the plate is not simply loaded or supported, it is often necessary to resort to other methods. Among the most common of these is to numerically model the membrane. This process, known as the method of finite differences, separates the plates into discrete points and analyzes the plate piecewise. This method is discussed in detail in Chapter 7.

ii) Lamb Waves

Engineers also utilize waves on plates as transducers. There are two kinds of waves that travel in plates. They are dilation waves, which involve changes in volume without rotation, and distortion waves, which do not change volume but instead result in rotation and shearing of a given material. In more common terminology, the dilation wave is referred to as a longitudinal wave, while the distortion wave is often called a transverse or shear wave. These waves travel

with a velocity that is material dependent and the respective velocities of each, c_1 and c_2, are related by the ratio:

$$\frac{c_2}{c_1} = \sqrt{\frac{(1-2\nu)}{2(1-\nu)}} \tag{6-21}$$

given that:

$$c_1 = \sqrt{\frac{\lambda + 2G}{\rho}} \tag{6-22a}$$

$$c_2 = \sqrt{\frac{G}{\rho}} \tag{6-22b}$$

where λ is Lamé's Parameter, defined by $\nu = \frac{\lambda}{2(\lambda + G)}$

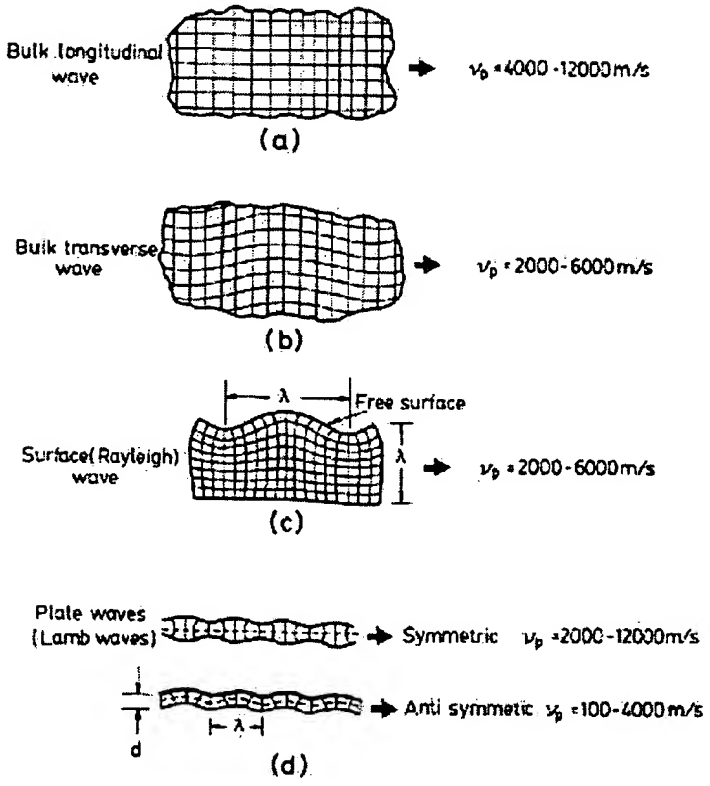


Figure 6-10: Wave propagation in solid media. (from [6])

In plates, these two waves interact in complex ways at the plate boundaries, which results in the formation of a plate wave, which is also called a lamb wave. These waves travel in either symmetric or anti symmetric modes as shown in Figure 6-10. The lowest order of these modes are very similar to surface acoustical waves, or SAW, that propagate along a semi-infinite medium. However, in thin plates, the lowest order symmetric mode is dispersionless and propagates much faster than a SAW on the same materials. The lowest order anti-symmetric mode, on the other hand, involves flexure and its wave velocity decreases monotonically to zero as the plate becomes infinitely thin.[51] Lamb waves travel with a phase and group velocity that is defined by:

$$v_{ps} = \sqrt{\frac{12D}{h^3\rho}} \quad v_{gs} = v_{ps} \tag{6-23a}$$

$$v_{pa} = \sqrt{wh\sqrt{\left(\frac{D}{h^3\rho}\right)}} \quad v_{ga} = 2v_{pa} \tag{6-23b}$$

where g and p represent group and phase, and s and a represent symmetric and anti-symmetric.

One of the interesting results of this analysis is that tension and wave velocity are coupled. If a small section of a plate is considered with dimensions, dx and dy, and a tension, T, in the x-direction, the out of plane force on the plate can be modeled by two forces, a stiffness and a tension:

$$dF_{zS} = -D\frac{\partial^4 w}{\partial x^4}dxdy \tag{6-24a}$$

$$dF_{zT} = T\frac{\partial^2 w}{\partial x^2}dxdy \tag{6-24b}$$

If these equations are combined and related to the acceleration of the membrane, the result is:

$$-D\frac{\partial^4 w}{\partial x^4}dxdy + T\frac{\partial^2 w}{\partial x^2}dxdy = Mdxdy\frac{\partial^2 w}{\partial t^2} \tag{6-24c}$$

where M is the mass per unit area of the membrane

If this equation is solved through separation of variables, one finds that the solution is:

$$w(x,t) = C_n e^{j(\omega_n t - k_n x)} \tag{6-25}$$

where

$$\omega_n = \frac{2\pi n}{P}\sqrt{\frac{1}{M}\left(t + \left[\frac{2\pi n}{P}\right]^2 D\right)}$$

P = period of the actuator that is oscillating the plate

n = integer representing the different modes of the device

This leads to the solution that the phase velocity is dependent on tension and mass:

$$v_p \approx \sqrt{\frac{T + \frac{2\pi n}{P} D}{M}} \tag{6-26}$$

As this analysis shows, the phase velocity is coupled to both the tension and mass density of the plate. This enables sensors that detect lamb waves to be sensitive to a wide range of different effects, with temperature and pressure changes being the more prevalent changes sensed. The advantage of using anti-symmetric lamb waves in sensors is that, on very thin plates, they have a phase velocity that is usually much slower than that of sound in most media. This allows these devices to transmit waves without dissipating large amounts of energy to the surrounding environment. In comparison to surface acoustic waves, which dissipate on the order of 1 dB per wavelength, lamb waves have extremely low loss mechanisms. In lamb waves, the disturbance to the surrounding medium only extends to a distance of $\lambda/2\pi$, which limits the acoustical energy loss. For an in-depth discussion of the physical properties of Lamb waves, Reference [115] treats the material more thoroughly.

iii) Modal Waves

While lamb waves have many applications in membranes, it is also useful to excite standing waves on plates. A standing wave, as opposed to a lamb wave, involves oscillations in fixed spots. These waves have maximum displacement at the resonant frequency of a device. On a resonating plate, there will be distinct spots called nodes, where vertical motion is essentially zero, and spots called anti-nodes, where oscillations are maximized. The analysis of standing waves begins with a dynamic version of Equation 6-18:

$$D\nabla^4 W = Mh\omega_n^2 W + N_x \frac{\partial^2 W}{\partial x^2} + N_{xy} \frac{\partial^2 W}{\partial x \partial y} + N_y \frac{\partial^2 W}{\partial y^2} \tag{6-27}$$

where

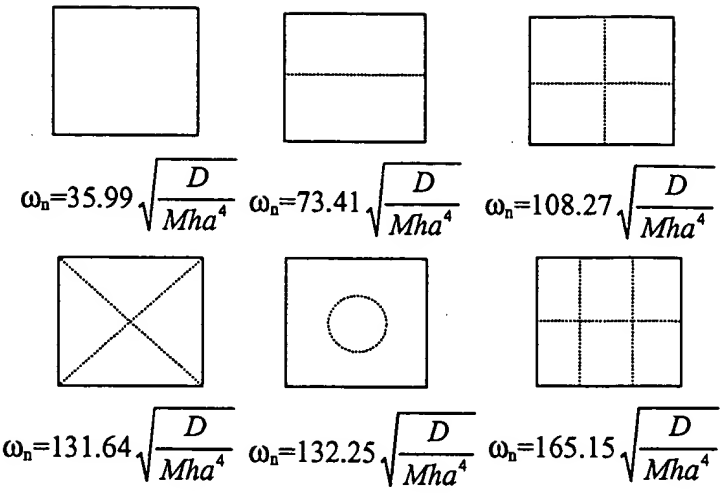


Figure 6-11: First six vibratory modes and resonant frequencies in a square plate with nodal lines shown.[54]

N_x, N_y = Normal loadings in the x and y directions

N_{xy} = shear loading

$W \equiv w = AW(xy)\cos(\omega_n t+\theta)$ where A is the oscillating amplitude

To solve this equation, the boundary conditions for W must be found that fit the end conditions. Since many square MEMS membranes are clamped on all sides, Figure 6-11 shows the modes and resonant frequencies of these structures. For other solutions to oscillating plate problems, Reference [54] offers an excellent analysis of plate mechanics.

B. Failure of Membranes

As membranes can be considered two dimensional equivalents of one dimensional beams, they have similar failure considerations. If Equation 6-18 is solved analytically, then the stresses on the plate are determined by the equations:

$$\sigma_x = -\frac{Ez}{1-\nu^2}\left(\frac{\partial^2 w}{\partial x^2}+\nu\frac{\partial^2 w}{\partial y^2}\right) \tag{6-28a}$$

$$\sigma_y = -\frac{Ez}{1-\nu^2}\left(\frac{\partial^2 w}{\partial y^2}+\nu\frac{\partial^2 w}{\partial x^2}\right) \tag{6-28b}$$

$$\tau_{xy} = \frac{Ez}{1+\nu}\frac{\partial^2 w}{\partial x \partial y} \tag{6-28c}$$

where z is the distance from the neutral axis of the plate

In order to determine the maximum stress in a plate, the following relationships are also useful:

$$M_x = -D\left(\frac{\partial^2 w}{\partial x^2} + \nu\frac{\partial^2 w}{\partial y^2}\right) \tag{6-29a}$$

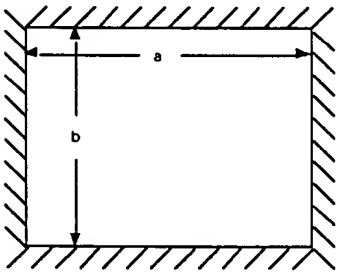
$$M_y = -D\left(\frac{\partial^2 w}{\partial y^2} + \nu\frac{\partial^2 w}{\partial x^2}\right) \tag{6-29b}$$

If the bending moments are known, the stress distribution can be calculated. Stress is zero along the neutral axis and rises linearly to a maximum at the surface. This maximum value is:

$$(\sigma_x)_{max} = \frac{6M_x}{h^2} \tag{6-30a}$$

$$(\sigma_y)_{max} = \frac{6M_y}{h^2} \tag{6-30b}$$

This analysis can be simplified for a membrane with a uniform loading, q, and fixed boundary conditions on all four sides. These structures will have the following stresses:



stress at center of long edge: $\sigma_{max} = \frac{-\beta_1 qb^2}{t^2}$ (6-31a)

stress at center: $\sigma = \frac{\beta_2 qb^2}{t^2}$ (6-31b)

The parameters α, β_1 and β_2 are functions of the plate geometry and boundary conditions, and may be determined from the table below:

a/b	1.0	1.2	1.4	1.6	1.8	2.0	∞
β_1	0.3078	0.3834	0.4356	0.4680	0.4872	0.4974	0.5000
β_2	0.1386	0.1794	0.2094	0.2286	0.2406	0.2472	0.2500
α	0.0138	0.0188	0.0226	0.0251	0.0267	0.0277	0.0284

Table 6-2: Plate coefficients (four sides fixed)[1].

Solutions for other geometries and loading conditions are also available; the reader is referred to References [58] and [59] for additional information.

Another area of concern in plate mechanics is the effects of internal stress upon deflection and strength. As discussed in Chapter 3, thin films can often have large residual stresses. As this stress is coupled to temperature, changes in temperature will also affect the output of many membrane based sensors.

A problem with using lamb wave oscillators is that their sensitivity is coupled to a number of different changes. Using these devices in space applications will be especially difficult due to the fact that they are natural thermocouples. For lamb wave oscillators to have a future in the aerospace industry, it must be proven that they are sufficiently decoupled from many common aerospace phenomena, such as large temperature and pressure changes, to be effective transducers.

C. Additional Reading

Timoshenko, S. and Woinowsky-Krieger, Theory of Plates and Shells: 2nd edition, New York: McGraw-Hill, 1959.

I. A. Viktorov, Rayleigh and Lamb Waves: Physical Theory and Applications, Plenum Press: New York, 1967.

[1] The discussion on stresses on a bounded rectangular plate is adapted from work done by K. Man at JPL.

III. Hinges

In MEMS, there is a need for devices that can produce out of plane motion without the limitations of a torsional spring. In these instances, hinges are often used.

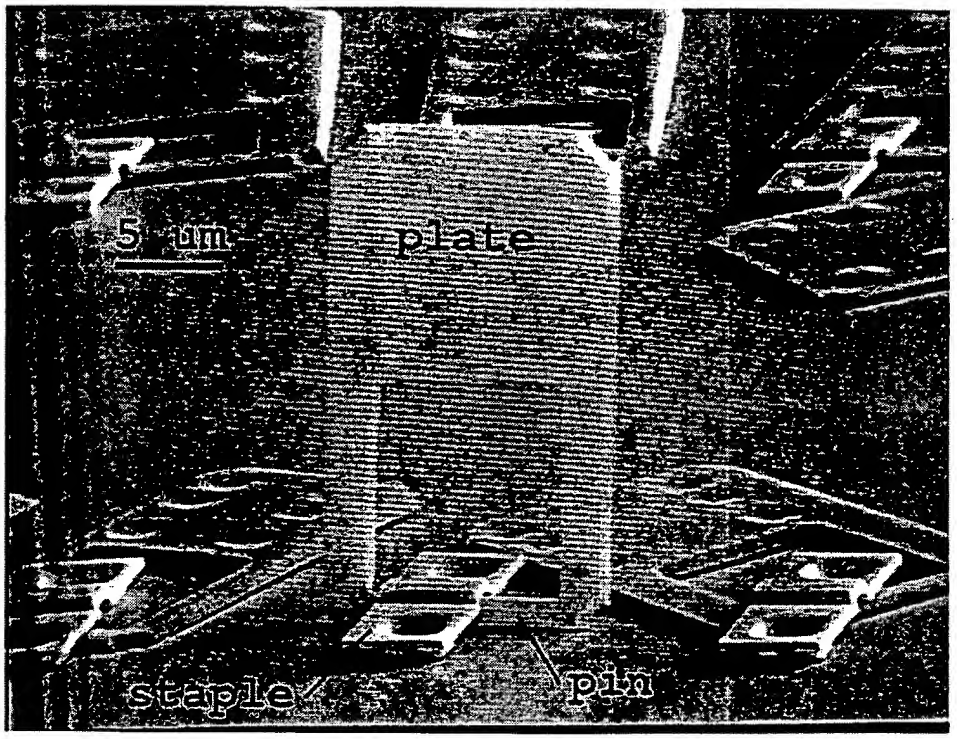


Figure 6-12: A typical hinge. (from [93])

In technical parlance, a hinge is an end condition that prevents translation of a structure, but allows free rotation. Often in MEMS, flexible structural beams that have hinged-boundary properties are called hinges. However, since the mechanical and reliability characteristics of these devices are similar to that of structural beams, there is no need to repeat that material here. It is enough to treat those devices as structural beams with narrow cross sections that create stress concentration at the interfaces with thicker beams.

This section will discuss surface micromachined non-planar hinges. These hinges can perform a multitude of tasks. One common implementation is to use hinges to hold structures, which were fabricated in a planar position, out of plane. Another common use of hinges is to bind non-planar structures together, as in the case of a cage or a box. This allows the fabrication of extremely high aspect ratio structures by common surface micromachining methods. This technology has enabled multitudes of new devices, such as optical devices and microgrippers. One of the biggest advantages of hinges is that they enable devices to be both

thermally and electrically removed from the substrate, which limits much of the noise common to planar sensors.[93]

A. Structural Analysis

A common hinge is depicted in Figure 6-12. These devices are simply constructed with only two parts. There are several types of hinges that were initially reported by Pister et. al. in [93], which are represented in Figure 6-13. 6-13a shows a substrate hinge, which is constructed out of a pin and a staple. The pin is a structural beam held down by the staple, which is a curved membrane. This hinge is often used to support non-planar structures and is fairly common in optical MEMS technologies. The other two hinges are called scissor hinges. They are constructed of interlocking beams, as shown, and usually have a wider range of motion. Scissor hinges are usually used to hinge released structures to each other. In many applications, hinges are not used to support large ranges of motion, but rather are used to support static structures. If the hinges are to be used in this static mode, a layer of material is deposited after assembly to bond the hinges into position.

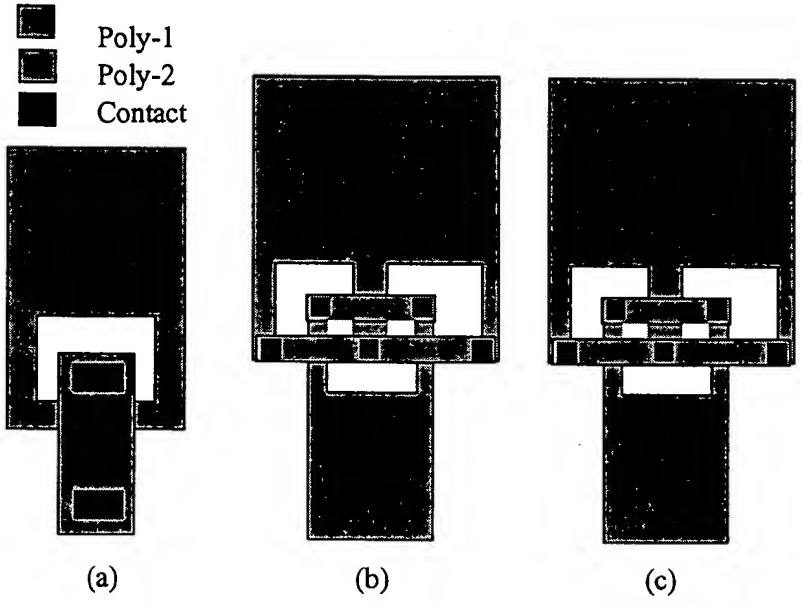


Figure 6-13 (a-c): Three basic hinges as presented by Pister et al.[93]

B. Reliability Concerns

One of the limitations in designing a substrate hinge is in designing the pin so that it is not wider than the allowable clearance of the staple. While this is an easy enough issue to address in design, it does provide some upper limits on the strength of the hinge. Scissor hinges, on the other hand, do not have limitations on the width of the beams. As a result scissor hinges are considered stronger structures than substrate hinges. If the hinge is anchored in place by the deposition of material after the device has been assembled, the adhesive strength of the deposited material will determine the hinge strength. Although Pister et al. reported a PECVD oxide layer that can withstand a torque of 10 nNm, this data would have to be independently determined by individual foundries.

Another area of concern in hinged devices is the issue of assembling hinged devices. While this is commonly done with micromanipulator stages in the laboratory setting, it may be extremely difficult to do this on a reproducible basis on a production line. As a result, membranes that are supported by hinges must be examined for damage caused in the assembly process.

C. Additional Reading

K. S. Pister, M. W. Judy, S. R. Burgett, and R. S. Fearing, "Microfabricated Hinges" *Sensors and Actuators A*, Vol. 33, pp. 249-256, 1992.

M. E. Motamedi, M. C. Wu and K. S. Pister, "Micro-opto-electro-mechanical Devices and On-chip Optical Processing" *Optical Engineering*, Vol. 36, No. 5, May 1997.

IV. Piezoresistive Transducers

Piezoresistivity is the property of a material whereby the bulk resistivity changes under the influence of a stress field. While all materials have varying degrees of piezoresistive responses, piezoresistors are commonly employed in semiconductor sensors because many semiconductor materials have large piezoresistive responses. The actual physics behind piezoresistive devices is slightly involved, but necessary to understanding its effects.

A. Formal Definition

To understand piezoresistivity, several other concepts must be first introduced. For a three-dimensional, anisotropic crystal, the electric field is related to a current by a three-by-three resistivity tensor given below.

$$\begin{bmatrix} \varepsilon_x \\ \varepsilon_y \\ \varepsilon_z \end{bmatrix} = \begin{bmatrix} \rho_1 & \rho_6 & \rho_5 \\ \rho_6 & \rho_2 & \rho_4 \\ \rho_3 & \rho_4 & \rho_3 \end{bmatrix} \cdot \begin{bmatrix} i_x \\ i_y \\ i_z \end{bmatrix} \tag{6-32}$$

If a Cartesian coordinate system is aligned to the <100> axis, then ρ_4, ρ_5, and ρ_6 become correlation coefficients, which relate the electric field in one axis to the current in a perpendicular direction. This leads to the result that in an isotropic conductor, such as unstressed silicon, $\rho_1=\rho_2=\rho_3=\rho$ and $\rho_4=\rho_5=\rho_6=0$. These values can be related to incremental changes in resistivity by the following equation:

$$\begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \rho_4 \\ \rho_5 \\ \rho_6 \end{bmatrix} = \begin{bmatrix} \rho \\ \rho \\ \rho \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} \Delta\rho_1 \\ \Delta\rho_2 \\ \Delta\rho_3 \\ \Delta\rho_4 \\ \Delta\rho_5 \\ \Delta\rho_6 \end{bmatrix} \tag{6-33}$$

To define the piezoresistivity, all one needs to do is to relate the fractional change in resistivity, $\Delta\rho_i/\rho$ to the stresses in the crystal. In order to do this fully, a 6x6 matrix must be defined. But for a crystal, this matrix will exhibit the same symmetries as the crystal lattice itself, which will obviate the need for many of the matrix coefficients. If the coefficients are defined as π_{ij}, a cubic crystal structure will only have three non-vanishing coefficients. For a silicon lattice, the matrix becomes:

$$\frac{1}{\rho}\begin{bmatrix}\Delta\rho_1\\ \Delta\rho_2\\ \Delta\rho_3\\ \Delta\rho_4\\ \Delta\rho_5\\ \Delta\rho_6\end{bmatrix}=\begin{bmatrix}\pi_{11} & \pi_{12} & \pi_{12} & 0 & 0 & 0\\ \pi_{12} & \pi_{11} & \pi_{12} & 0 & 0 & 0\\ \pi_{12} & \pi_{12} & \pi_{11} & 0 & 0 & 0\\ 0 & 0 & 0 & \pi_{44} & 0 & 0\\ 0 & 0 & 0 & 0 & \pi_{44} & 0\\ 0 & 0 & 0 & 0 & 0 & \pi_{44}\end{bmatrix}\begin{bmatrix}\sigma_x\\ \sigma_y\\ \sigma_z\\ \tau_{xy}\\ \tau_{xz}\\ \tau_{yz}\end{bmatrix} \qquad (6\text{-}34)$$

If all these equations are combined, it is possible to get an expression for the electric field as a function of stress[1]:

$$E_x = \rho i_x+\rho\pi_{11}\sigma_x i_x+\rho\pi_{12}(\sigma_y+\sigma_z)i_x+\rho\pi_{44}(i_y\tau_{yz}+i_z\tau_{xz}) \qquad (6\text{-}35a)$$

$$E_y = \rho i_y+\rho\pi_{11}\sigma_y i_y+\rho\pi_{12}(\sigma_x+\sigma_z)i_y+\rho\pi_{44}(i_x\tau_{yz}+i_z\tau_{xy}) \qquad (6\text{-}35b)$$

$$E_z = \rho i_z+\rho\pi_{11}\sigma_z i_z+\rho\pi_{12}(\sigma_x+\sigma_y)i_z+\rho\pi_{44}(i_x\tau_{xz}+i_y\tau_{xy}) \qquad (6\text{-}35c)$$

These equations clearly show that there is a direct relationship between stress and resistivity. It is also important to note that materials with small piezoresistive coefficients will have more limited responses than those with larger coefficients. One interesting aspect of Equation 6-34 is that it closely resembles Hooke's law. In fact, this tensor is relating a resistivity strain, instead of a mechanical strain, to the stresses upon a material. As such, there is also an orientation dependence on the piezoresistive coefficients similar to the one discussed for the elastic moduli in Section 3-IA. For an in-depth discussion of this material, Chapter 4 in Reference [6] offers a more complete mathematical description of piezoresistivity.

B. Piezoresistive Sensors

Devices that utilize the piezoresistive effect are designed so that mechanical stress occurs simultaneously with an event to be measured and that the stress is proportional to the magnitude of the event. Currently there are two kinds of piezoresistive sensors made. Membrane sensors are manufactured to measure pressure and flow while cantilever beams sensors are made for accelerometers.

Membrane sensors are usually designed as a thin single crystal silicon plates supported by a thick ring. Usually a piezoresistor is built into the edge of the device to utilize stress concentration. When the membrane deforms under an externally applied load, there will be

[1] This discussion assumes an infinite bulk lattice. For finite crystals there is a small correction factor due to dimensional changes.[6]

stress on the piezoresistor. On cantilever beam sensors, the piezoresistor is, for similar reasons, placed on the surface of the beam near its support.

If it is assumed that the mechanical stress over the resistor is constant, the change in resistance can be given as:

$$\frac{\Delta R}{R} = \sigma_l \pi_l + \sigma_t \pi_t \tag{6-35d}$$

where

σ_l, σ_t = longitudinal and transverse stresses

π_l, π_t = longitudinal and transverse piezoresistance coefficients

For a resistor made of p-type material this expression reduces to

$$\frac{\Delta R}{R} = \frac{\pi_{44}}{2}(\sigma_l - \sigma_t) \tag{6-36}$$

For n-type resistors the expression becomes

$$\frac{\Delta R}{R} = \frac{\pi_{11} + \pi_{12}}{2}(\sigma_l + \sigma_t) \tag{6-37}$$

One important feature of these equations is that, due to the fact that they assume uniform stress fields, they are only valid for resistor sizes much smaller than the membrane or beam size.

Usually piezoresistors are configured in a Wheatstone bridge. Two resistors are placed to measure stress parallel to current flow, while two are placed to measure stress perpendicular to current flow. This arrangement works so that any decrease in resistance from tensile stress is balanced by a corresponding increase in resistance for compressive stress. This has the effect of creating a differential output of opposite signs on each side of the bridge. The total voltage change is defined as

$$\Delta V = \frac{\Delta R}{R} V_b \tag{6-38}$$

where V_b is the voltage applied to the bridge, as shown in Figure 6-14.

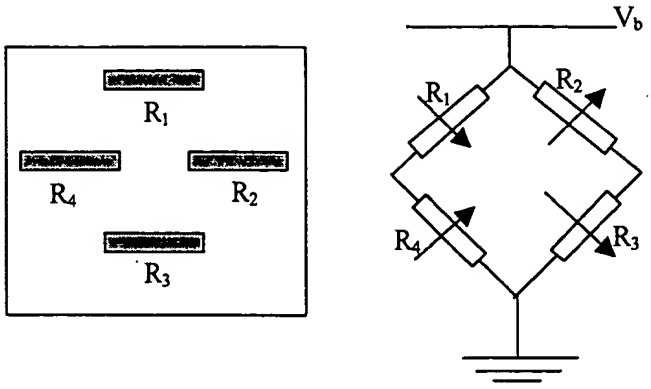


Figure 6-14 (a,b): Schematic representation of the position of four piezoresistors on a membrane (left) and accompanying circuit diagram (right). The arrows represent resistance changes when the membrane defects in the –z direction.[6]

C. Reliability Issues

One of the problems that will be encountered in using piezoresistors in high-rel applications is that they exhibit a temperature dependence. If the relationship between the piezoresistive coefficients and temperature is plotted, it becomes apparent that there is a roughly linear relationship between log(π) and log(T). For a generalized description, the piezoresistive coefficient can be determined as a function of both doping concentration, N, and temperature, T, by:

$$\pi(N,T) = \pi_0 P(N,T) \tag{6-39}$$

where π_0 is the low-doped room temperature piezoresistive coefficient.

P(N,T) is offered graphically in Figure 6-15. As can be clearly seen, at low doping concentrations, there is better sensitivity but a greater temperature dependence. As doping concentrations becomes greater than 10^{20} atoms/cm^3, the temperature dependence becomes indiscernible, but sensitivity decreases greatly. For space applications, with the great thermal ranges usually required, these devices will almost certainly have to balance the sensitivity requirements with the temperature dependence.

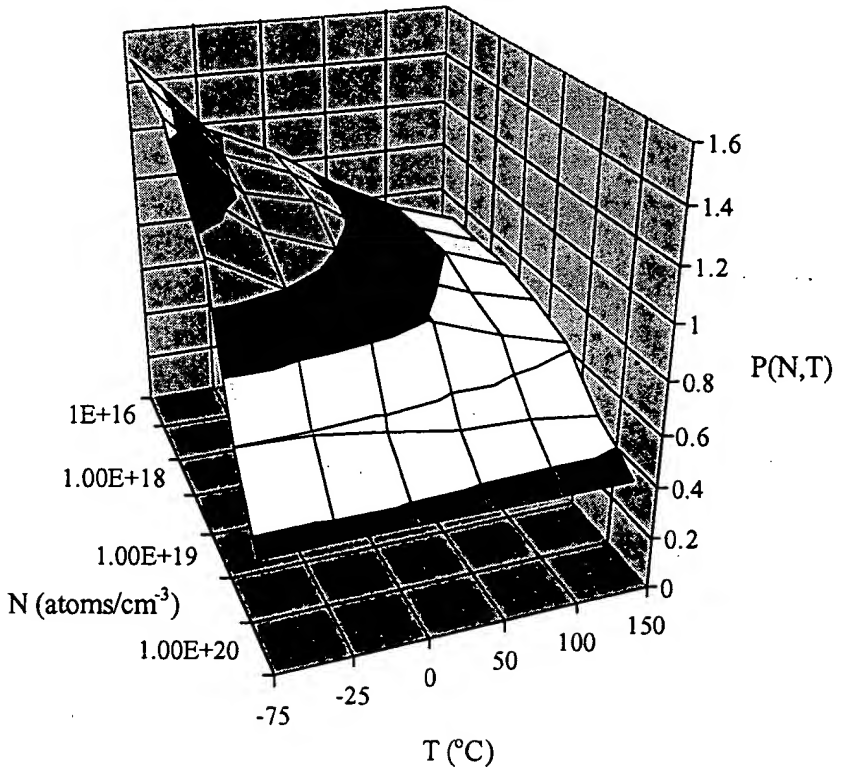


Figure 6-15: P(N,T) for n silicon. As can be seen, P(N,T) converges to a nearly uniform, albeit smaller, value versus temperature at doping levels above 10^{20} atoms/cm^3.

D. Additional Reading

Sze, S. M. ed., Semiconductor Sensors, Wiley Inter-Science, New York, 1994, Cp. 4.

Y. Kanda, "A Graphical Representation of the Piezoresistive Coefficients in Silicon", *IEEE Transactions on Electron Devices*, Vol. ED-29, No. 1, January, 1982.

V. Tunneling Tips

Electron tunneling is a concept that was developed in this century as an outcropping of quantum theory. Tunneling developed from the study of the energy of an electron in a confined space. Basic quantum theory entails Schrödinger's equation, which describes a particle's wave function, ψ, by the relation:

$$\nabla^2\Psi = -\frac{2m}{\hbar^2}(E-U)\Psi \tag{6-40}$$

where

h = Planck's constant of 6.62617×10^{-34} J-s

E = total energy of the particle

U = potential energy of the particle

m = mass of the particle

The wave function of a particle is usually not as informative as the value of the wave function multiplied by its complex conjugate, $\psi^2=\psi\times\psi^*$. This value represents the probability that a particle will be in a given point in space. If this equation is solved for the case of an electron that is in a one-dimensional energy well of width a, bounded by two infinite potential energy barriers, as depicted in Figure 6-16, the solution to Schrödinger's equation yields:

$$\psi^2(x)=0 \ \{x<0\}$$

$$\Psi^2(x) = \left(\frac{2}{a}\right)\sin^2\left(\frac{n\pi x}{a}\right) \ \{0<x<a\} \tag{6-41}$$

$$\psi^2(x)=0 \ \{x>a\}$$

where n is an integer.

This equation shows that the electron is bound by the infinite barriers, and can never escape from the well it is in. However if the barrier at a is replaced with a barrier of finite energy and width, then solving Schrödinger's equation shows that the electrons will actually tunnel through the barrier and there will be a non-zero probability that there will be an electron on the other side of the barrier. This means that it is possible for electrons to actually pass through areas in which, according to classical physics, they do not have the energy to penetrate. It is this phenomenon, called electron tunneling, that is utilized to produce tunneling tip sensors.

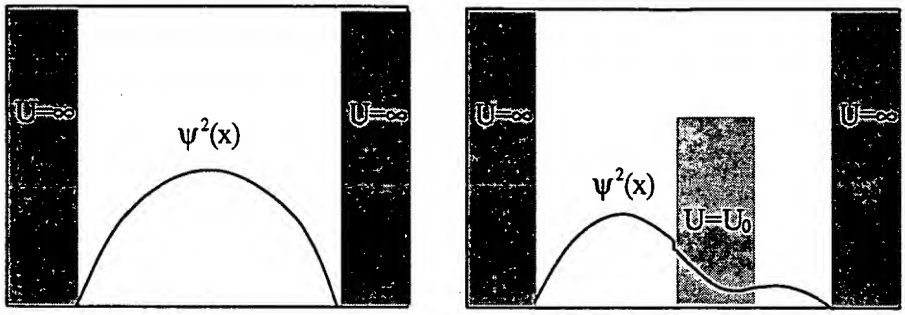


Figure 6-16: Probability distribution of an electron trapped in a well. The figure on the left shows an electron bounded by two infinite walls while the figure on the right shows an electron bounded by two finite walls with a finite energy barrier, which is greater than the energy of the electron, in the middle. As can be seen, there is a finite probability that the electron will penetrate the barrier and be on the other side. This diagram roughly corresponds to the device in Figure 6-18.

Tunneling tips are small pointed tips, shown below, that were initially developed for use in electron microscopy. They have since been adopted by the MEMS community because the tunneling effect is an extremely accurate way to measure displacements caused by external effects.

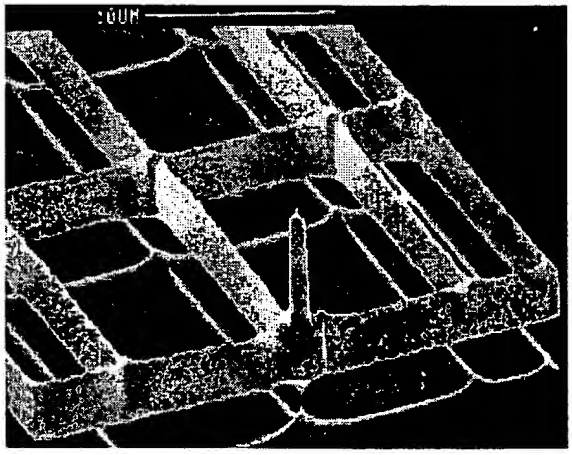


Figure 6-17: Tunneling tip on SCS beams. (from [154])

A. Physical Description

It has been shown that the current caused by tunneling across a narrow barrier is given by:

$$I \propto Ve^{(-\alpha x\sqrt{\phi})} \tag{6-42}$$

where

V = bias voltage on the tip

$\alpha = 1.025\ \text{Å}^{-1}\text{ev}^{-1/2}$

ϕ = the effective energy height of the tunneling barrier = $-\frac{1}{\alpha I}\frac{\partial I}{\partial V}\frac{\partial V}{\partial x}$

x = the physical width of the energy barrier

This means that an electronic circuit capable of detecting a 1% variation in a 1 nA current from a 100 MΩ source would be able to detect deflections on the order of 0.003 Å. For this reason tunneling tips have started to be developed for use in high data storage applications and high sensitivity accelerometers.

Typically tunneling sensors are designed by suspending a mass above the tunneling sensor. An external force, which can be anything from infrared radiation to acceleration, pushes the mass downwards, which increases the tunneling current and becomes a measurable event, as shown in the diagram below.

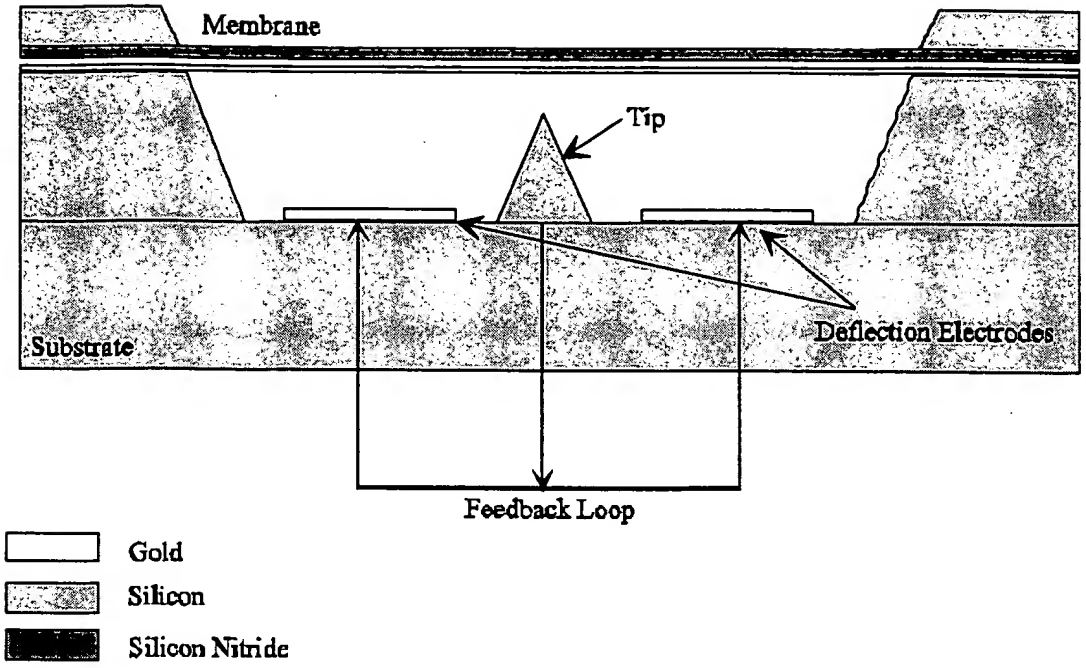


Figure 6-18: Typical layout for a tunneling sensor.

While tunneling sensors are among the most accurate sensors available in MEMS technology, they do suffer from a number of reliability problems.[41,43]

B. Reliability Concerns

One of the great difficulties in making tunneling accelerometers is in fabricating the devices. The tunneling tips need to be made of a conductive surface that does not react with the air. While conventional tunneling tips, since they operate under an Ultra-High Vacuum, can be made from a multitude of metals, microsensor tunneling tips are much more limited in the materials that can be used. Gold has been found to be useful in the production of sensors, but it is difficult to create good adhesion between a gold tip and an insulating substrate, which is usually SiO_2. This created a need for multiple layers of adhesive materials, which create processing problems and reduce yield.[42] While the fabrication of these devices is certainly not an impossible task, they do suffer from low yield rates.

Another problem that is more difficult to handle stems from the fact that the tunneling effect is highly displacement sensitive. Since, in order to get a tunneling effect, the tip must usually have a bias voltage, which is typically under 1 volt, and be within nearly ten Ångstroms of the moving mass, contact between tip and mass is unavoidable. This contact must be accounted for in design of mechanical system and circuit. On the mechanical side, it is important to place one of the electrodes on a compliant support to limit the force to the tip during the

inevitable contact. The circuit must also limit the current during contact. If both of these precautions are taken, the danger of tip crashes will be mitigated. For devices designed with these techniques, crashes have occurred at low frequency operation for months with no detectable change in operational characteristics. However, any tunneling tip designs need to have these issues thoroughly addressed for high-rel applications.

These devices, due to their extreme sensitivity, are also susceptible to thermal noise and mechanical vibration. It is typically these effects that limit device sensitivity and they need to be addressed for any tunneling sensor. There is also electrical noise in the measurements of the tunneling output. In many devices, the electrical noise spectrum exhibits a 1/f dependence. This noise creates an error in measurements that scales on the order of 10^{-3} Å/Hz^{½}. While this noise is fairly insignificant for many applications, it does provide a limit on the actual sensitivity of the device.[41]

C. Additional Reading

T. W. Kenny, W. J. Kaiser, H. K. Rockstad, J. K. Reynolds, J. A. Podosek, and E. C. Vote, "Wide-Bandwidth Electromechanical Actuators for Tunneling Displacement Transducers" *Journal of Microelectromechanical Systems*, Vol. 3. No. 3, September 1994.

VI. Electrostatic Actuators and Transducers

A. Parallel Plate Capacitors

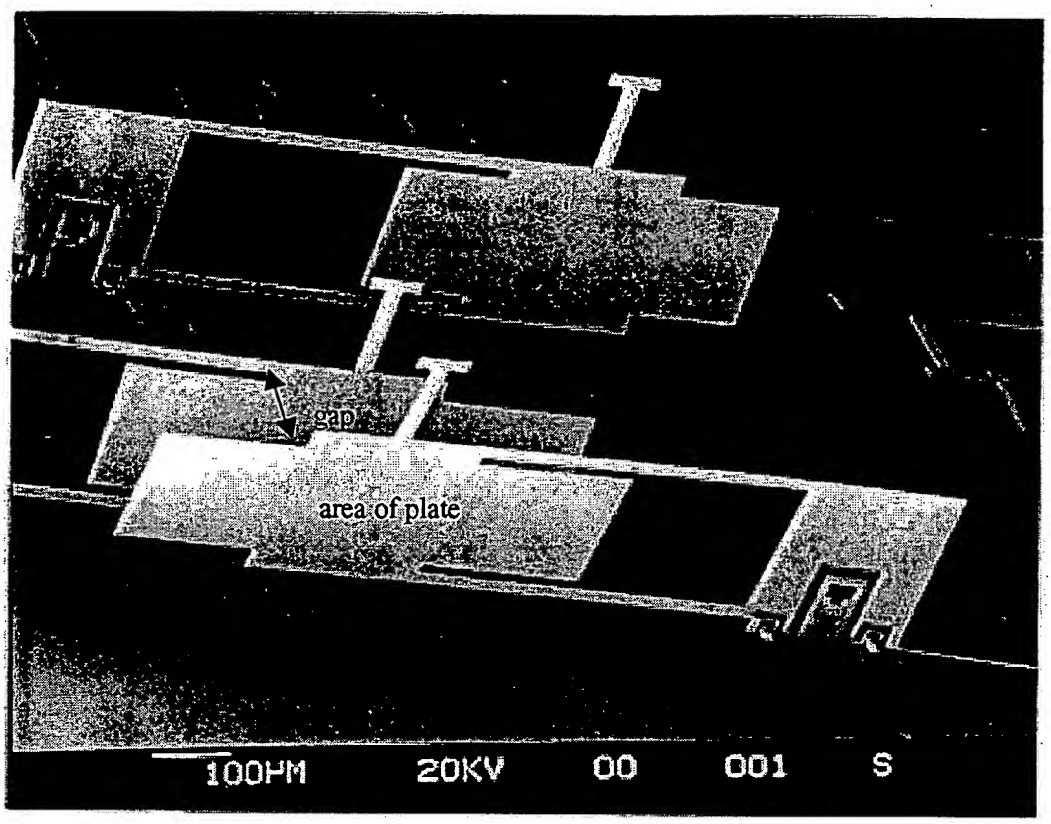


Figure 6-19: Closeup view of parallel plate capacitor with the area and gap labeled. (from[155])

Capacitors have been developed and understood for as long as any electronic device. They are both simple to construct and understand. As such, capacitors are fundamental to many devices and have been used extensively in the microelectronics industry. It has been estimated that a modern microprocessor has anywhere between seven and ten million recognizable capacitors in its design. A capacitor is simply two conductive objects separated by some distance, d, which store electrical energy by attracting and repelling free electrons within the conductors. For MEMS, a prevalent capacitors design consists of two parallel plates, as shown in Figure 6-19.

i) Electrical and Mechanical Analysis

The capacitance of an object is defined as the amount of electric charge that it can store per voltage. The common mathematical expression of a capacitor's ability to store energy or charge is related by the expressions:

$$q = CV \tag{6-43a}$$

$$U = \tfrac{1}{2}CV^2 \tag{6-43b}$$

where

q = electric charge

V = voltage

C = capacitance

U = energy

For capacitors constructed of two parallel plates a useful relationship has been derived that:

$$C = \varepsilon \frac{A}{d} \tag{6-44}$$

where

A = area of one of the plates

d = distance between the plates

ε = the permittivity of the material between the two plates

One reason that capacitors have become prevalent in MEMS is that capacitance is a function of distance. This means that a change in distance will result in a change in capacitance, which is a measurable event. If a capacitor immersed in air is assumed to have one fixed plate and one plate that displaces a distance, x, from the rest, the capacitance can be rewritten as:

$$C = \varepsilon_0 \frac{A}{(d - x)} \tag{6-45}$$

where

ε_0 = permittivity of a free space ($8.85*10^{-12}$ F/m)

If a sensor is to be fabricated out of parallel plate capacitor, there is a simple method to electrically detect a change in capacitance. Since a current, I, is related to the charge on a capacitor, Q_c, by the equation:

$$i_c = \frac{dQ_c}{dt} \tag{6-46}$$

Then using the circuit below,

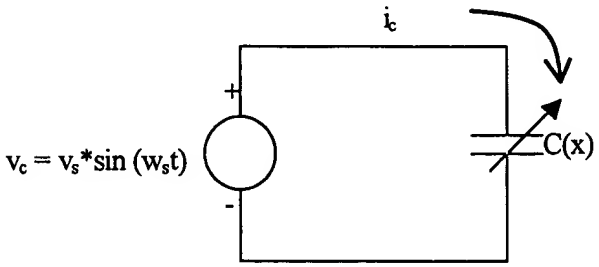


Figure 6-20: Basic circuit for detecting changes in capacitance.

a change in current will be approximately related to a change in displacement by the relationship[1]:

$$\Delta i_c(\Delta x) \approx \omega_s v_s C(d)\left[\frac{\Delta x}{d}\right] \tag{6-47}$$

Another useful feature of parallel plate capacitors is the fact that they can be made into actuators. Given that a force, F, is related to potential energy by the equation:

$$F = -\frac{\partial U}{\partial x} \tag{6-48}$$

it is possible to derive the relationship for parallel plate capacitors that:

$$F = \frac{\varepsilon_0 A V^2}{2d^2\left(1-\frac{x}{d}\right)^2} \tag{6-49}$$

This shows that an applied voltage will exert a force on the capacitor plates. It is this electrostatic force that is used to make actuators out of parallel plate devices.

---

[1] This is a linearized result that only applies for $\Delta x << d$.

ii) Limitations of Parallel Plate Capacitors

While parallel plate capacitors have good actuation and sensing abilities, they have some severe limitations. The greatest drawback to using these devices is that they are non-linear. While they can be treated as linear for small displacements, for large motions, parallel plate capacitors clearly exhibit non-linear behavior.

One of the dangers in these devices is the potential of the plates touching. As discussed in Chapter 3, when two metal surfaces come into contact, adhesive forces exert a strong bond that usually causes failure. This problem is especially prevalent in parallel plate devices because of the non-linear force that increases quadratically with distance. A common design rule used is that, if $\Delta x \geq 1/3$ d, the device will usually have sufficient force to transverse d. To prevent this, parallel plate devices must be designed to displace much less than this amount.[10]

Parallel plate capacitors could also be susceptible to electrostatic discharge. An ESD would have a similar effect as applying a delta function to the device. If the voltage spike is large enough, it could induce stiction by bringing the plates into contact. Unfortunately the scant research into ESD in MEMS has not provided any concrete data on the effects of ESD on parallel plate capacitors and these theories have not been experimentally verified.

iii) Additional Reading

W. S. Trimmer, K. J. Gabriel, and R. Mahadevan, "Silicon Electrostatic Motors", *Transducers '87, The 4th International Conference of Solid-State Sensors and Actuators*, pp. 857-860, June 1987.

W. S. Trimmer and K. J. Gabriel, "Design Considerations for a Practical Electrostatic Micro-Motor", *Sensors and Actuators*, Vol. 11, pp. 189-206, 1987.

B. Comb Drives

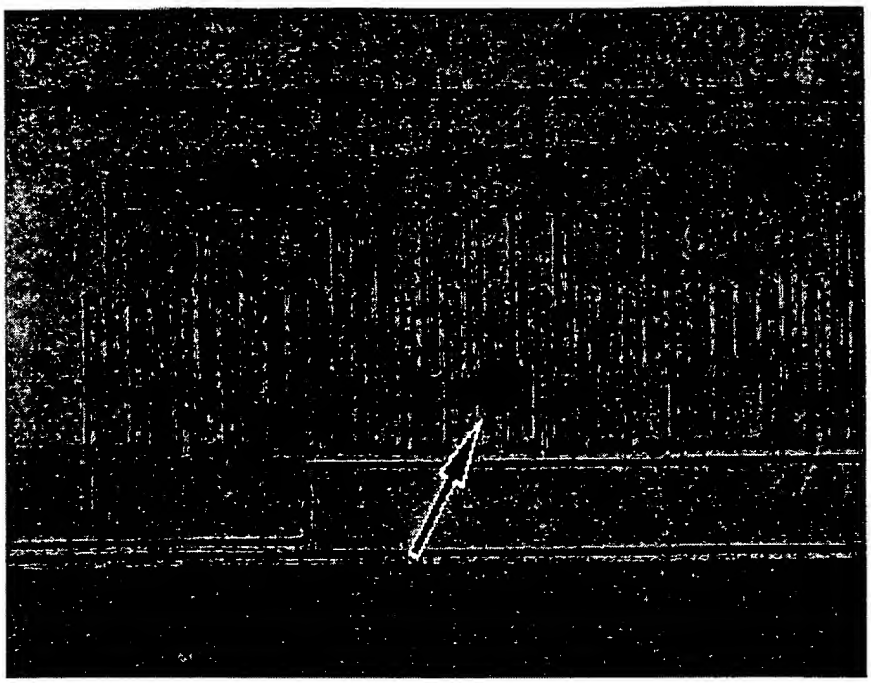


Figure 6-21: A standard comb drive. The arrow points to a particulate that landed on one of the electrodes.

Due to the problems discussed with making electrostatic parallel plate actuators, there have been attempts made to make devices that utilize electrostatics to produce motion while eliminating the relationship between force and distance. The most common device made to accomplish this is called a comb drive because of its overall comb-like appearance, as seen in Figure 6-16. Comb drives operate by using fringing fields to pull one set of the drive into the other. Actuation occurs in one dimension only and the equations of motion are derived in the next section.

i) Mechanical and Electrical Analysis

Essentially comb drives are composed of multiple structural beams and, as such, they are not difficult to analyze. As shown in the picture, there are two sets of interdigitated electrodes. Generally one of them is fixed while the other is mobile. The force generated by a voltage is described by:

$$F = \frac{\varepsilon_0 b V^2}{d} \tag{6-50}$$

where b is the vertical height of the cantilever beams.

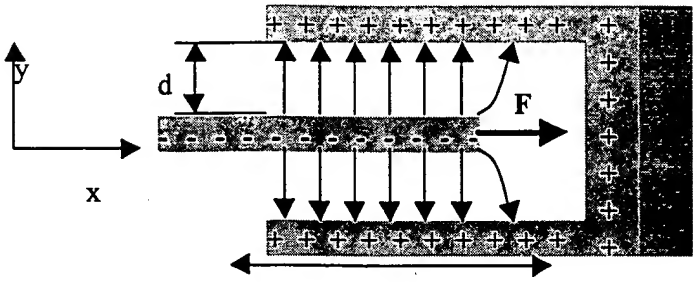


Figure 6-22: Overhead view of basic unit of a comb drive.

This equation shows that the force produced by this device is independent of the displacement of the middle electrode. As a result, comb drives are among the more common actuators in MEMS. The other fact that makes comb drives appealing is that n comb drives produce a force equal to n×F. This simple scaling makes it possible to produce macroscopic forces on the MEMS scale. It is not unreasonable to manufacture an array of comb drives to produce upwards of 30 Newtons with as little as 10 volts applied to the electrodes.[10] While these aspects of comb drives are quite appealing, there are also some serious limitations to their performance.

ii) Limitations on Design

One of the problems with comb drives illustrated by Figure 6-22 is that there is one electrode. For argument's sake, assume it to be negatively charged, surrounded by two positively charged electrodes. While the forces on the electrode are balanced by the equal spacing of the two gaps, clearly any perturbation of the center electrode will cause an offset of the forces and pull the electrode to one side or another. In engineering terms, the comb drive is an inherently unstable system. To combat this problem, several steps must be taken in the design to insure that the device does not fail at the first vibration. Since stability problems are problems in energy storage, the key to designing a stable comb drive is to design it to store more energy in the y direction than in the x direction. To make $U_y >> U_x$, the following condition must be met:

$$\tfrac{1}{2}k_y d^2 >> \tfrac{1}{2}k_x l_p^2 \qquad (6\text{-}51a)$$

where

k_y = spring constant in the y direction

k_x = spring constant in the x direction

l_p = The length of the electrode

which means that:

$$k_y >> \frac{l_p^2}{d^2} k_x \qquad (6\text{-}51b)$$

This is illustrated in Figure 6-23.

While this is hardly an exact answer to the problem of stability, this solution does lend it some formalism. For high-rel applications, a large safety factor will have to be included to guarantee that the comb drive does not fail due to surface contact. If a comb drive is not entirely stable, it is possible for the drive to not completely fail but instead enter into chaotic oscillatory modes. While this phenomenon has applications in encryption, it is usually an undesirable event.

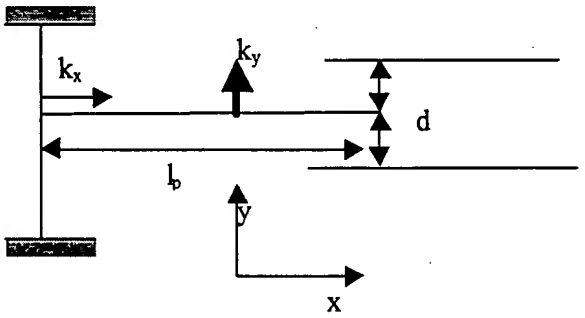


Figure 6-23: Overhead view of comb drive structure with spring constants and dimensions labeled.

Another limitation on comb drives is that they are typically limited in the amount of work they can do. With a maximum displacement of l_p and, with l_p being kept to a minimum due to Equation 6-51, comb drives are just not capable of producing large scale motion. While this

problem can be designed around to a degree, it is nevertheless a serious limitation to comb drive usage.

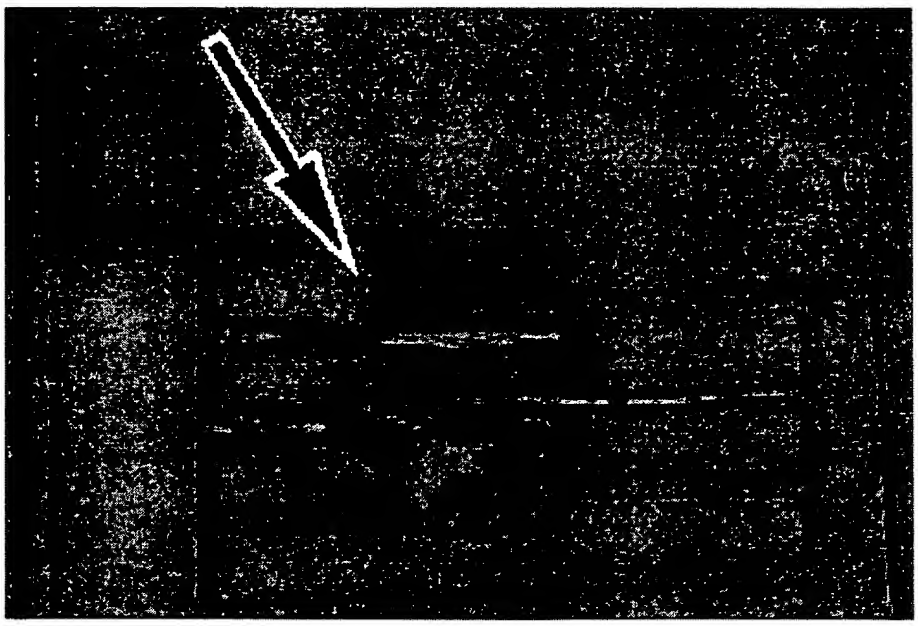


Figure 6-24: An example of the damage that results when two comb drive electrodes come into contact and short out the device.

An area that is also important to consider in comb drive operation is the effects of parasitic capacitance with the substrate. Since the comb drive is a fairly large conductive surface suspended over another large conductive surface, there is a considerable parasitic effect between the substrate and the drive. While this effect can be used to produce out of plane torsional microactuators, it is often an undesirable side effect of the comb drive design. It is possible to have such a large parasitic motion that the comb drive will actually touch the substrate, which will lead to the adhesion and possibly shorting problems. In a sound design, the comb drive should be far enough removed from the substrate that the parasitic capacitance will not cause stiction.

Particulates can also be problematic in comb drives. Conductive dust particles can electrically connect parts of a comb, which will short them out, producing catastrophic current flows. Comb drives could also be susceptible to electrostatic discharge. An ESD would have a similar effect as applying a delta function to the device. If the voltage spike is large enough, it could induce stiction by bringing the plates into contact. Unfortunately there is no published information on the effects of ESD on comb drives, which means that evaluating the ESD tolerance of a design is not yet possible.

iii) Additional Reading

W. C. Tang, T.-C. H. Nguyen, and R. T. Howe, "Laterally Driven Polysilicon Resonant Microstructures." *Proceedings of IEEE Microelectromechanical Systems*, February 1989.

W. C. Tang, T.-C. H. Nguyen, M. W. Judy and R. T. Howe, "Electrostatic-comb Drive of Lateral Polysilicon Resonators" *Transducers '89, Proceedings of the 5th International Conference on Solid-State Sensors and Actuators and Eurosensors III*, Vol. 2, pp. 328-331, June 1990.

C. Micromotors

A major area of research in MEMS in the past decade has been into the design and fabrication of micromotors. There are multiple kinds of micromotors being designed today. While most of these devices are electrostatically driven by side electrodes, as illustrated by Figure 6-25, there are also a number of other designs being implemented. However, for the sake of brevity, this discussion will be limited to electrostatically driven micromotors. The section on reliability will have implications to less conventional micromotors.

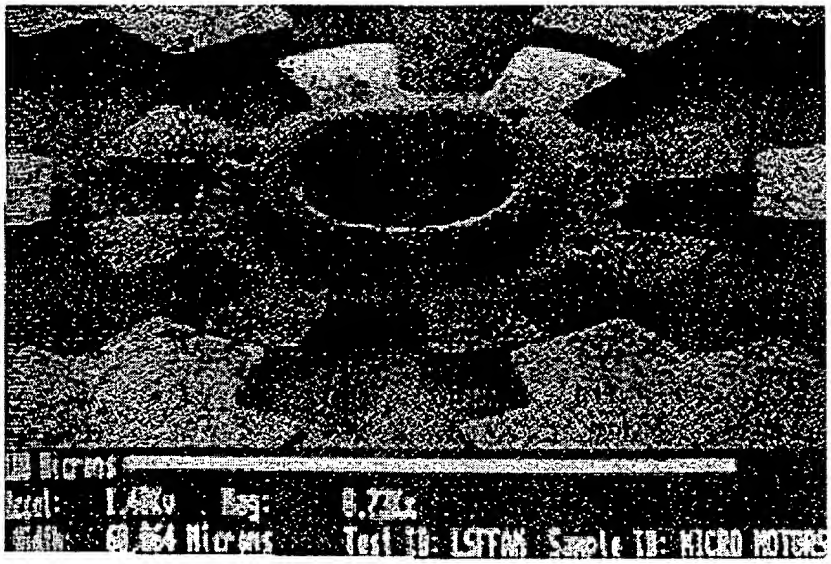


Figure 6-25: Electrostatically driven micromotor. (from [22])

i) Electrostatic Motor Analysis

Electrostatic micromotors utilize variable capacitance in a fashion somewhat similar to other electrostatic devices previously discussed. The main difference is that micromotors are typically driven by several different sets of drives, or stators, that are switched on and off to produce a torque. This torque is a function of the rotation angle of the drive:[22]

$$T(\theta) = \tfrac{1}{2} V^2 \frac{\partial C(\theta)}{\partial \theta} \tag{6-52}$$

Micromotors have a natural operating frequency, which is determined by the magnitude of torque applied to the motor. For a motor similar to the one depicted above, this frequency is:

$$f_N(V_p) = 1.5\left(\frac{V_p}{100}\right)^2 kHz \tag{6-53}$$

where V_p is the phase voltage applied to the stators. This leads to an equation for the maximum rotational speed of the motor, ω_{max}:

$$\omega_{max} = \frac{240 f_N}{n} rpm \tag{6-54}$$

where n is the number of steps per revolution determined by[1]:

$$n = \frac{1}{\left(\frac{1}{n_s} - \frac{1}{n_r}\right)} \tag{6-55}$$

where n_s and n_r are the respective numbers of stators and rotors.

These motors can be either used as microstepper motors or can be operated as a continuously rotating motor. While the actual design and fabrication of these devices varies depending upon application, most electrostatic motors are governed by the above physical laws.[22,23]

ii) Harmonic Motors

Harmonic motors are a kind of motor that utilizes the rolling motion of two bodies with different circumferences. Typically a motor contains a cylindrical hole, with a slightly smaller cylindrical rotor. These motors have a tribological interest as they utilize rolling instead of sliding friction, which eliminates many of the wear concerns common to other motors. Furthermore, since the stators and rotors are designed to touch in these devices, large amounts of force can be generated.

Using one of the original harmonic motors as an example, there are some basic facts of motor operation that are fairly common to all harmonic motors. If the stators are cycled at a frequency of ω_s, then it is possible to determine the steady state frequency of the harmonic motor, ω_r:[87]

$$\omega_r = \omega_s \left(\frac{R_s}{R_r} - 1\right) \tag{6-56}$$

where R_s and R_r are dimensions defined in Figure 6-24.

---

[1] This derivation assumes a drive signal described by Y. –C. Tai et al. in [22].

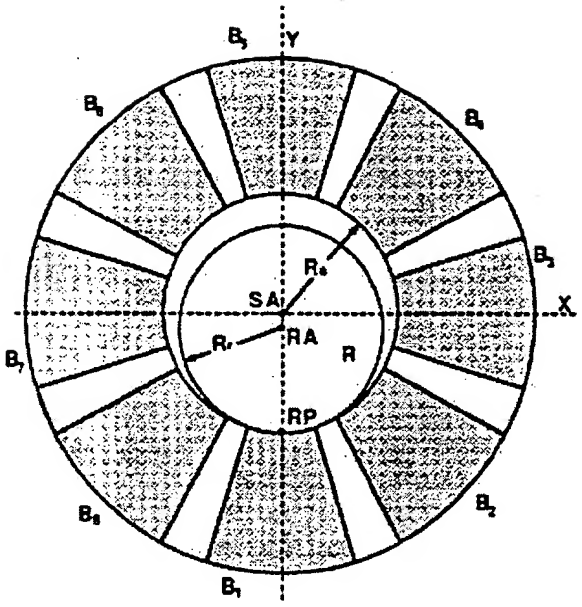


Figure 6-26: A basic harmonic motor layout. (from [88])

The unsymmetrical nature of the roll causes the rotor to effectively wobble inside the stator. One reliability concern of wobble motors is that the direct contact of the stator and rotor raises questions about stiction.

iii) Microbearing Reliability Concerns

The biggest reliability concern for micromotors is the motor's connection to the substrate. Since the rotors must be supported on bearings, there is concern about the long term reliability of these bearings. It has been shown that, over time, there will be wear on the bearings caused by the frictional forces from the substrate. The wear on the bearings will increase the frictional force on the bearing, requiring higher driving voltages, which will further increase wear. This positive feedback loop will quickly lead to total device degradation. The only effective method to mitigate wear on bearings is to select wear- resistant materials.

In most macroscopic devices, liquid lubricants are used to prevent direct metallic contact. However, in MEMS, it is felt that liquid lubricants will not be in general use due to the fact that viscous friction forces are large compared to other frictional forces on micrometer scales.[19] While recent research has raised the possibility of using gas phase lubricants to reduce wear, these solutions involve extremely high temperature operations and are unlikely to be practical.[151] Instead, bearings are usually operated dry at ambient temperature, with direct contact made between structures. There have been studies conducted on these

conditions and it has been found that different materials respond, as would be expected, in distinct ways to wear.

Several studies[19,153] indicate that single crystal silicon is moderately well suited as a bearing material. During initial burn-in, the rough points in a bearing will fracture off, leaving a smoothed surface that shows little wear over time. As a result, the wear on silicon decreases with time. Polysilicon shows moderate wear properties and is a suitable, although not ideal, bearing material. As would be expected, Si_3N_4 and SiO_2 exhibit poor bearing characteristics. They show linear wear that leads to total failure. Diamond-like carbon is a material that has shown promise as a bearing material and may eventually be used as a coating on many SCS and polysilicon structures that have large contact stresses. While the exact characteristics of bearing wear are dependent upon the materials involved and the environment in which they slide, a good rule of thumb is that wear is minimized by using dissimilar hard materials

Another problem associated with wear on microbearings is that, for electrostatically driven structures, forces on the device will be a function of the device height, and thus bearing height. Many devices place rotors above stators in order to have a non planar component of force to partially levitate the rotor. As the bearings wear, the electrode distance would decrease, which would cause forces to increase quadratically. This increases wear and alters drive performance. This behavior accelerates failure and leads to total device collapse. Ultimately, the contact morphologies of microbearings are the limiting factor in micromotor performance and, as such, they need to be well understood for high-rel applications.

iv) Additional Reading

Long-Sheng Fan, Yu-Chong Tai, and Richard S. Muller, "IC-Processed Electrostatic Micromotors." *IEEE International Electronic Devices Meeting*, December 1988.

Yu-Chong Tai, Long-Sheng Fan, and Richard S. Muller, "IC-Processed Micromotors: Design, Technology, and Testing" *Proceedings of IEEE Microelectromechanical Systems*, February 1989.

U. Beerschwinger, D. Mathieson, R. L. Reuben, and S. J. Yang. "A study of Wear on MEMS Contact morphologies" *Journal of Micromechanics and Microengineering*, 7 September 1994.

S. C. Jacobsen, R. H. Price, J. E. Wood, T. H. Rytting, and M. Rafaelof, "The Wobble Motor: An Electrostatic, Planetary Armature, Microactuator" *Proceedings of IEEE Microelectromechanical Systems*, pp. 17-24, February 1989.

VII. Magnetic Actuators

Magnetic actuators are a class of devices that, as their name implies, utilize magnetic fields to provide force. While creating magnetic fields on semiconductor devices is intrinsically more difficult than creating electric fields, the potential benefits of magnetic actuators has spurred the development of these devices. Due to the physics of magnetics, magnetic devices are capable of producing greater forces than electrostatics. Combined with the ability to apply force through a conductive medium, such as electrolytic fluids, these factors make magnetic actuators a promising field within MEMS.[15]

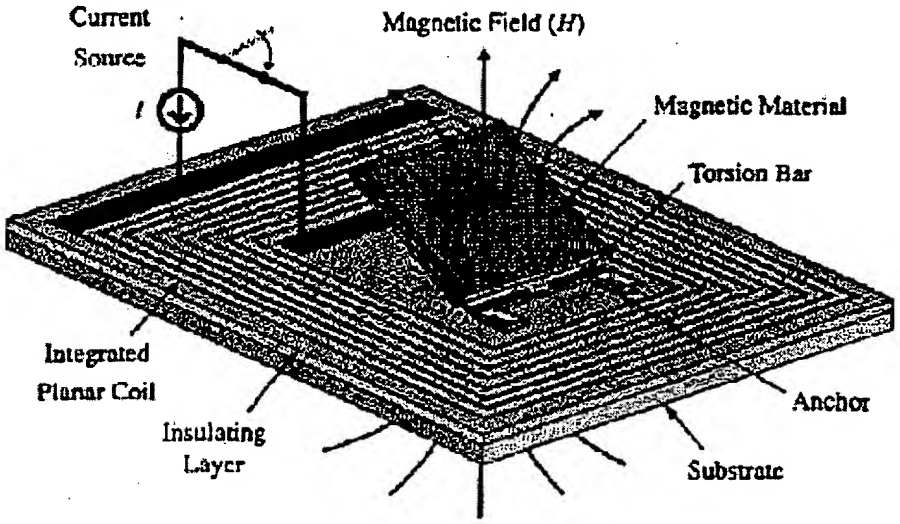


Figure 6-27: A magnetic actuator designed by Judy et al. (from [15])

There are several common methods employed in MEMS to make magnetic actuators. A magnetic field can be described from Maxwell's equations:

$$\nabla \times B = \left(\frac{4\pi}{c}\right) J + \left(\frac{1}{c}\right)\frac{\partial E}{\partial t} \tag{6-57a}$$

$$\nabla \cdot B = 0 \tag{6-57b}$$

where

B =magnetic field

c = the speed of light

E = electric field

J = current density

These equations show that a magnetic field can be produced either by a constant current or by a time varying electric field. While there are multiple methods to create one of these two effects, the simplest way to create an actuating magnetic field is through a loop of wire, as shown in Figure 6-20. For these actuating fields to create motion, the actuator requires a structure that is influenced by the magnetic field. The two most common structures that will be actuated in a magnetic field are those with current loops on their surfaces and those coated with magnetic films, such as ferromagnetic and diamagnetic materials. In either case, the interaction of the two magnetic fields creates an actuating force that then moves the structure.

As with many of the technologies within MEMS, magnetic actuators are being developed in a myriad of ways by different groups around the world. Since magnetic actuators have not become as standardized in MEMS as electrostatic actuators, it is difficult to discuss a typical magnetic actuator. For this discussion, a simple actuator developed by Judy et al. will be discussed to give an example of the forces at work within a magnetic device. This device, which is shown in Figure 6-27, uses a ferromagnetic plate influenced by a loop of wire integrated into the substrate.

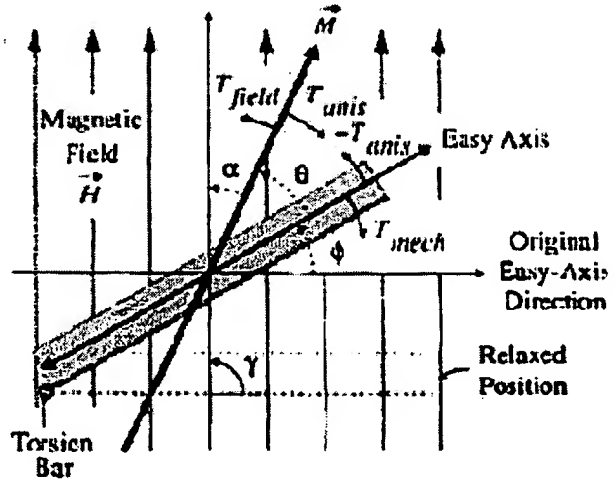


Figure 6-28: Diagram of forces acting on soft magnetic plate. (from [15])

A. Mechanical and Electrical Analysis

For this non-planar magnetic microactuator, the rotational deflection is a function of the magnetic field, H, and the stiffness of the torsion bar, k_ϕ. Assuming that the magnetic field remains perpendicular to the original orientation of the plate, the torque produced by the magnetic field, T_{field} is defined by:

$$T_{field} = V_{mag} M (H_{dc} + H_{ac} \sin(2\pi f t) \cos\phi) \tag{6-58}$$

where

M = net magnetization vector

V_{mag} = magnetic volume

H_{dc}, H_{ac} = magnetic field from respective dc and ac sources

f = ac current frequency

ϕ = angle plate rotates from rest.

To determine the actual mechanical response of the plate, it is necessary to use the dynamic torsional model of

$$J\ddot{\phi} + C_\phi \dot{\phi} + k_\phi \phi = T_{field} \tag{6-59}$$

where

$C\phi$ = dampening coefficient = $\sqrt{\frac{Jk_\phi}{Q}}$

J = polar moment of inertia of the magnetic plate

Q = quality factor

To determine ϕ as a function of time, Equation 6-59 must be solved. This can be done either analytically or numerically, although the analytic solutions might be difficult to obtain. In order to do either, M must be defined, which is done below:

$$M = \frac{\mu_0 (\pm H_c + H_a)}{N_M} \tag{6-60}$$

where

H_c = coercive field of the magnetic material under torque

H_a = applied field ~ $H*\cos(90-\phi)$

N_M = shape-anisotropy coefficient of the plate

μ_0 = permeability of free space ($4\pi*10^{-9}$ H/m)

The net result of the forces is that the plate will oscillate in the time varying magnetic field. While these equations give a basic description of the motion of a ferromagnetic plate in a magnetic field, other magnetic materials will have different responses to similar fields. Due to the fact that the exact internal effects of magnetic devices extend beyond the scope of this guideline, an interested reader should examine the references listed at the end of the section to gain a more detailed description of the physics of these devices.

B. Reliability Concerns

There are some basic problems with using magnetic forces in MEMS that need to be understood. An inherent drawback to magnetic devices is that they scale poorly into the micro domain. In order to scale a device effectively, certain quantities must be held constant while the physical dimensions of a device shrink. If three basic quantities of current density, heat flux, and temperature rise are considered to be held constant during scaling, it becomes apparent that there are serious limitations to micro-magnetic actuators.

If current density is held constant while scaling a device, then a wire with an order of magnitude drop in the cross sectional area will have an order of magnitude drop in current. While this results in a constant heat generation per unit volume, it will, for a wire- generated field operating on a permanent magnet, result in a drop in force of three orders of magnitude. While this loss can be slightly offset by the fact that smaller systems are better at conducting heat away, clearly this method of scaling severely limits the effectiveness of micro-magnetic actuators. If instead the heat flux per unit of surface area for a wire is constant during scaling, current density scales according to the inverse square root of the change in heat flux, so that, for an order of magnitude drop in cross sectional area, there will be an increase in force on the order of 2.5 orders of magnitude. This scaling is limited by the maximum allowable temperature on a device and is also not a desirable method to increase force as dimensions decrease. If the system is scaled to limit the temperature difference, it is possible to have a two order of magnitude increase in force. However, this comes at the expense of an increase in current density, which makes the device much less efficient, which may not be acceptable for many applications. Thus, magnetic devices cannot scale into the micro domain without sacrificing either force, operating temperature, or efficiency, which is a serious limitation.[55] As a result, any magnetic actuator design must be verified to insure that it meets the design requirements without unacceptable temperature dissipation or power losses.

Another issue to consider in using magnetic actuators aboard spacecraft is the presence of spurious magnetic fields. Since modern spacecraft employ an array of electronic devices that create magnetic fields, these devices can be unintentionally actuated by nearby devices. A good model of the magnetic fields in a spacecraft, which is usually developed by the systems engineers, will determine the risk level for parasitic actuation.

C. Additional Reading

R. E. Pelrine, "Room Temperature, Open-Loop Levitation of Microdevices Using Diamagnetic Materials", *Proceedings of IEEE Microelectromechanical Systems*, pp. 34-37, February 1990.

R. E. Pelrine and I. Busch-Vishniac "Magnetically Levitated Micromachines", *IEEE Micro Robots and Teleoperators*, November 1987.

J. W. Judy and R. S. Muller, "Magnetic Microactuation of Torsional Polysilicon Structures," *Sensors and Actuators A, Physical*, Vol. A53, Nos. 1-3, pp. 392-397, 1996.

J. W. Judy and R. S. Muller, "Magnetically Actuated, Addressable Microstructures", *Journal of Microelectromechanical Systems*, Vol. 6, No. 3, September 1997.

VIII. Thermal Actuators

Thermal actuators are a class of devices that utilize heating to produce forces and deflections. These devices operate through the use of heat transport to rapidly change a device's temperature. Since MEMS devices operate on such small scales, it is possible to create devices with quick response times, as heat transport occurs in scales often measured in microseconds. While some objections may be raised to the power dissipation implicit with these devices, they offer a simpler alternative to many electrostatically and magnetically driven devices.

A. Bimetallic Strips

The most prevalent thermally actuated devices in MEMS are structures constructed out of layered materials. These thermal actuators utilize the bimetallic effect found in common household thermometers. There have been a number of arguments made for the advantages of bimetallic actuators. Since there is a direct coupling between dissipated power and beam deflection, the actuators can operate at low voltage levels. Combined with the ability to produce a force that is independent of displacement, thermal actuators have piqued serious interest and have been developed independently by a number of researchers.[150]

B. Mechanical Analysis

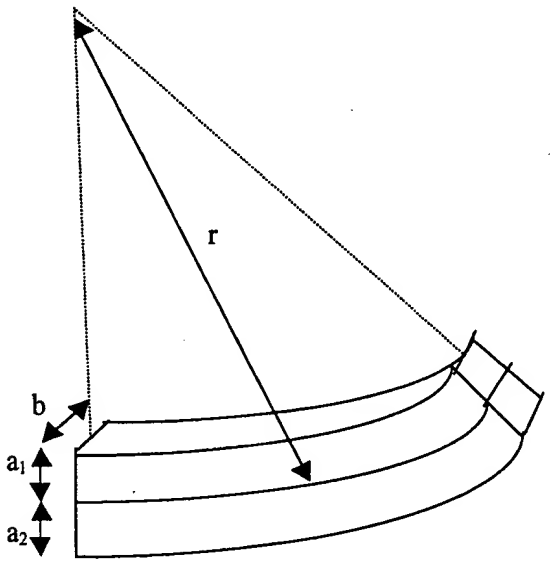


Figure 6-29: Side view of two metallic strips a and b bending due to temperature stress.

In these devices, two materials, often Si-SiO_2 or Si-Si_3N_4 are sandwiched together. As the device heats up, the differing changes in length caused by mismatched thermal expansion coefficient create stresses at the metallurgical junction, which bends the device. Figure 6-29 illustrates a bimetallic strip made of two cantilevered beams of lengths l that are exposed to a temperature change, ΔT.

The conversion factor γ, relates a temperature change, ΔT, with a deflection at the end of the cantilever, d, by:

$$d = \gamma \Delta T \tag{6-61}$$

If a uniform heat distribution is assumed with the beams, then the conversion factor can be approximated for $l << r$ as

$$\gamma = \frac{l^2}{2r\Delta T} \tag{6-62}$$

where r is the radius of curvature of the beam:

$$r = \frac{2}{3}\frac{\frac{7}{4}(a_1 + a_2)^2 - 2t_1t_2 + \frac{E_1b_1a_1^3}{E_2b_2a_2} + \frac{E_2b_2a_2^3}{E_1b_1a_1}}{(\alpha_1 - \alpha_2)\Delta T(a_1 + a_2)} \tag{6-63}$$

where α is the linear coefficient of thermal expansion.

A closer mathematical analysis of this equation will show that it is minimized when the cantilever beams have identical values of b and a. For these beams, the deflection is given by:

$$r = \frac{a}{3(\alpha_1 - \alpha_2)\Delta T}\left(5 + \frac{1+\chi^2}{\chi}\right) \tag{6-64}$$

where χ is the ratio $\frac{E_1}{E_2}$.

Since the absolute width of a structure does not influence bending, the smallest radius of curvature, and thus the greatest deflection, will occur if $(\alpha_1-\alpha_2)$ and l are maximized, while a is minimized. The simplest method to accomplish this comes through altering l and a, as there are only a few materials, and accompanying thermal expansion coefficients, to choose from in the semiconductor industry.[149]

As this derivation shows, the displacement of the strip is directly related to a change in temperature. There are several ways to induce temperature changes. One common method is

to apply an electric current through the beam. The power dissipated by the current flowing through a resistor will produce a ΔT. The amount of energy needed to raise temperature is determined by the heat capacity of the cantilever beams. In static operation, it is desirable to reduce all the dimensions of the actuator so that there is less thermal loss. In dynamic operation, as is the case for a high frequency switch, it is more useful to have larger surface areas with a greater heat exchange and a corollary increase in switching rate.[52]

C. Shape Memory Alloys

Shape memory alloy, or SMA, actuators are variants upon thermal actuators that use the shape memory alloy effect, which was first discovered in 1938 by Alden Greninger and V.G. Mooradian.[56] Materials that experience the SMA effect undergo reversible phase transformations. Below some critical temperature, the material is in the martensite phase and will easily deform. Above this temperature, the material changes to the austenite phase and begins to exert strong forces trying to restore to its original shape.

In the early 1960s, two researchers at the Naval Ordnance Laboratory discovered that the alloy NiTi can have a phase transition that is a function of alloying content and varies anywhere from –50 to 166°C. Since this material, nicknamed nitinol, for Nickel Titanium Naval Ordnance Laboratory, has superior mechanical properties, it is the material of choice in modern SMA research. While there are a number of applications of SMA materials, they all share some basic commonalities.[145]

A typical SMA uses a nitinol wire connected to a heater which, as previously discussed, can easily be a current flowing through a resistor. The materials properties of nitinol have been investigated and show the expected temperature dependence.

Properties	NiTi at 23 °C	NiTi at 110 °C
Young's Modulus	33.4 GPa	34.9 GPa
Linear Strain (at 10N)	1.6%	1.56%
Ductile Yield	20-30%	19-30%
Tensile strength	1.05 GPa	.962-.1.7 GPa
Resistivity	4400Ω/cm	4400Ω/cm
Linear coefficient of expansion	1.5×10^{-3} °C^{-1}	$.34\times10^{-3}$ °C^{-1}

Table 6-3: Properties of nitinol at different temperatures.[57]

The actual design of shape memory alloys varies wildly with applications. Studies have shown that nitinol springs can develop stresses in excess of 200 MPa. While these forces are impressive, SMA are not common MEMS devices and due to concerns discussed in Section C, may not see integration into space environments.

D. Reliability Issues for Thermal Actuators

One problem with making bimetallic thermal actuators is that they induce large stresses in the devices. These stresses can cause serious problems for long term reliability. It is not entirely clear how the interfaces that silicon forms with other materials will behave under repeated stressing and unstressing. Since these thermal actuators operate by stresses materials interfaces, there is an increased chance of fracture at this interface, which can lead to delamination. There is also an issue of long term thermal fatigue. Some of these devices are heated to above 800 °C and cooled to ambient temperature within the span of several tens of microseconds,[137] which could cause significant fatigue. The long term effects of this cycling is an issue that needs to be addressed in high-rel thermal actuators.

Thermal actuators are also frequency limited. The response of the actuator is governed by the time it takes for heat to convect and radiate away from a device. If the quantity 1/f is less than the time it takes a device to dissipate heat, oscillatory behavior will effectively stop, as illustrated in Figure 6-30. This causes some interesting concerns for the space environment, as there is no convective heat transfer in a vacuum. Thus the time required to dissipate heat should be significantly slower than it is in terrestrial applications.

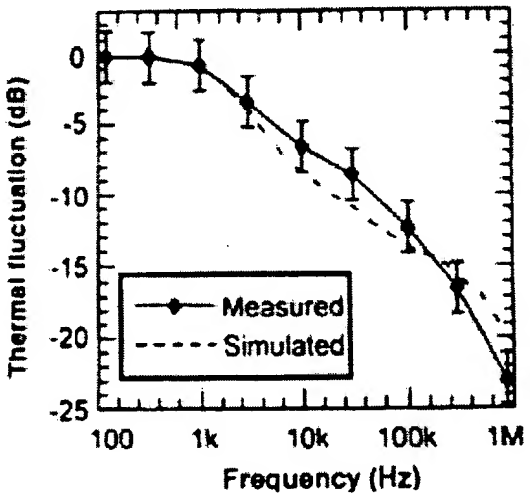


Figure 6-30: Mechanical response as a function of frequency for a thermal actuator. (from [117])

The time, t, required for a body to cool can be modeled by summing the sources of heat energy and equating them to the sinks. For a body radiating heat and having no convective transfer, this results in the equation:[156]

$$mc\frac{dT}{dt} = \sigma_b \varepsilon_m A\left(T^4 - T_0^4\right) \qquad (6\text{-}65)$$

σ_b = Stefan-Boltzmann constant (5.67×10-8 W/m^2-K^4)

ε_m = emissivity of the material

A = the surface area of the material

T_0 = initial temperature

By solving this equation, it is possible to get a limit on the time it takes for a radiative mass to cool, which indicates the frequency limit of a thermal actuator operating in the vacuum of space.

Shape memory alloys have problems unique to their structures. Since these alloys are usually made of ductile materials, they experience wear and fatigue at much faster rates than brittle materials. While they can withstand stress in the range of a 1 GPa, the lifetime and reliability of these devices at these stresses is unsuitable for long-term operation. If high reliability and millions of temperature cycles are desired, then nitinol should be stretched from its memory state only a few percent and should not exceed a couple hundred MPa of stress.

Another concern with these devices in space applications is inadvertent heating. Since most spacecraft experience wide temperature swings from periods of full solar exposure to eclipse, these devices will either have to have active on-chip thermal control or be in a well thermally regulated part of the spacecraft. Otherwise there could be disastrous implications in using thermal actuators in the space environment.

E. Additional Reading

W. Riethmüller and W Benecke, "Thermally Excited Silicon Microactuators" *IEEE Transactions on Electron Devices*, Vol. 35, No. 6, June 1988.

P. A. Neukomm, H. P. Bornhauser, T. Hochuli, R. Paravicini, and G. Schwarz, "Characteristics of Thin-wire Shape Memory Actuators" *Transducers '89, Proceedings of the 5th International Conference on Solid-State Sensors and Actuators*, Vol. 2, pp. 247-252, June 1990.

IX. Piezoelectric Actuators

Piezoelectric materials exhibit motion under an applied electric field. The piezoelectric effect has been well researched and understood for many years and the MEMS community has used piezoelectrics to build devices that produce strong forces with small actuation distances.

A. The Piezoelectric Effect

Piezoelectricity determines the distribution of the electric polarization and demonstrates how a piezoelectric field reacts to an electric stress by emitting depolarization waves.[6] This polarization field is linearly related to mechanical strain in certain types of crystals, such as quartz and GaAs. When the crystal is in equilibrium, strain is balanced by internal polarization force. However, when equilibrium is offset by external mechanical stress or by an external electric field, the emitting depolarization field will create a force to restore internal equilibrium. As a result, an externally produced electric field will cause a displacement and an externally produced mechanical stress will create an electric field.

Since the piezoelectric effect couples mechanical and electrical fields effectively, it has been researched in a multitude of materials. In 1910, Voigt showed that there were 32 classes of crystals that exhibited piezoelectric properties, and he measured coupling coefficients for these. In MEMS, the most common materials used are crystalline SiO_2 (Quartz), ZnO, AlN, and PZT.

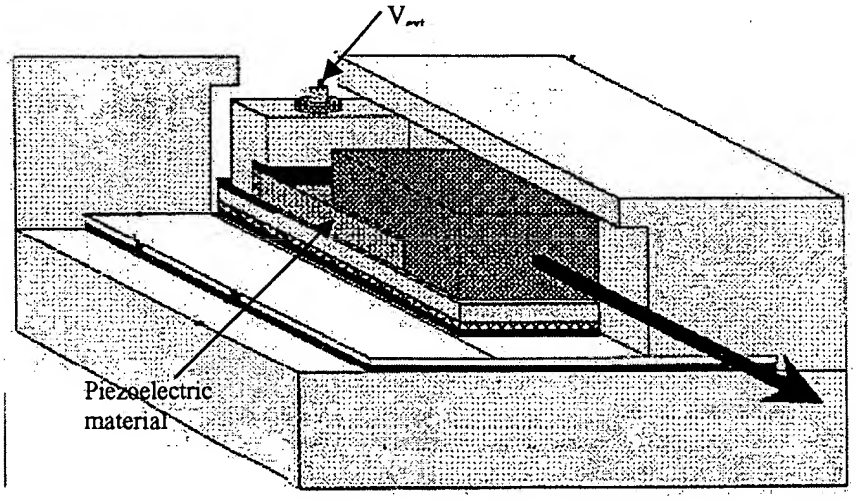


Figure 6-31: Diagram of a piezoelectric transducer (after [106]). An applied voltage causes the piezoelectric material to expand, which drives the structure.

B. Piezoelectric Devices

Piezoelectric devices can be constructed out of a number of different structures. A common implementation of piezoelectrics is to produce a piezoelectric mass and connect electrical leads to it, as shown in Figure 6-25. The extension and contraction of the piezoelectric bar is governed by the equation:

$$D = e_\varepsilon E + e\varepsilon \tag{6-66a}$$

$$T = c_E \varepsilon + eE \tag{6-66b}$$

where

E = the electric field

D = electric displacement at equilibrium

e_ε = dielectric constant at zero strain

e = is the piezoelectric stress constant

ε = mechanical strain

T =externally applied stress

c_E = elastic stiffness at equilibrium

So, clearly, the mechanical displacement in this structure is coupled to the applied electric field. One limitation of piezoelectric devices is that the actuation distance is usually small. Since piezoelectric devices operate by inducing a strain in a crystal, it would be extremely unusual to displace a piezoelectric device more than a few percent of its total length.

Piezoelectric devices are also commonly used as sensors. Since piezoelectric materials are electromechanically coupled, these devices can be used in much the same manner as piezoresistive elements, with strain being converted into a change in current instead of a change in resistivity.

C. Reliability Issues

Piezoelectric devices operate by inducing stress in a material. As such, they should be treated as structural devices and be analyzed for stress distributions to prevent fracture. The difficulty with structural analysis in piezoelectric devices stems from the fact that piezoelectric materials have unusual crystal structures. Quartz, for example, is a rhombohedral structure with nine individual elastic constants. For these materials Young's modulus and Poisson's ratio will

exhibit less symmetry than cubic crystals. While the analyses for these materials is not intractable, it is not as straightforward as it is for cubic crystals.

Piezoelectric devices also generate a considerable heat. Since these devices are subjected to mechanical stresses and have significant electrical losses, there is heat transfer across a device. For a rectangular piezoelectric actuator driven at a frequency f, there will be a change in temperature, ΔT, determined by:[116]

$$\Delta T = \frac{ufv_e}{k(T)A} \tag{6-67}$$

where

$$k(T) = \sigma_b \varepsilon_m \left(T^2 + T_0^2\right)\left(T + T_0\right) + \bar{h}_c$$

$\bar{h}_c$ = the average convective heat transfer coefficient (6-30 W/m^2-K in air)

u = loss of the material per cycle

A = surface area of the piezoelectric actuator

v_e = effective volume of the piezoelectric actuator (volume of the material not at equilibrium)

This heat production will stress a material and will also limit the performance of a device. Piezoelectricity is also, like piezoresistivity, temperature sensitive. As such, the heat generation must be considered in determining the reliability characteristics of a piezoelectric device.[116]

D. Additional Reading

J. W. Judy, D. L. Polla, and W. P. Robbins, "Experimental Model and IC-Process Design of a Nanometer Linear Piezoelectric Stepper Motor" *Microstructures, Sensors and Actuators*, DSC-Vol. 19, pp. 11-17, November, 1990.

K. Ikuta, S. Aritomi, T. Kabashima, "Tiny Silent Linear Cybernetic Actuator Driven by Piezoelectric Device with Electromagnetic Clamp", *IEEE Proceedings Microelectro-mechanical Systems*, pp. 232-237, February 1992.